Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

On the existence of stationary optimal policies for the average cost control problem of linear systems with abstract state-feedback

Alessandro N. Vargas and João B. R. do Val

***Abstract*— This paper establishes conditions for the existence of optimal stationary policies for a class of long-run average cost control problems. The discrete-time system is assumed to be linear with respect to the state but the controls take an abstract state-feedback structure. The derived approach may be used to represent systems where the state is observed by the controller only through some specially structures output (no history is employed). It is shown that, if there exists an optimal-abstract policy for the discounted-cost problem, and such a policy generates an autonomous system with uniform exponential decay, then there exists an optimal stationary policy for the average cost problem. Notions of controllability and observability of linear time-varying systems are imposed.**

***Index Terms*— discrete-time systems, feedback control, controllability, observability, linear-quadratic problems, optimal stochastic control, Markov processes.**

I. Introduction

Consider a discrete-time linear system modeled by the following evolution difference equation:

$$x_{k+1} = Ax_k + Bu_k + Ew_k, \quad \forall k \geq 0, \quad x_0 \in \mathbb{R}^n, \tag{1}$$

where x_k, u_k, and w_k evolve respectively, in $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^q$ and they represent the system state, control variable, and additive noisy input, in this order. As usual, the matrices A, B and E, of respective dimensions, are given.

There exist many systems for which the controller does not have complete information on the system state x_k. In the stochastic control literature the optimal control problem is studied, taking into account the past history of an observation process. This optimal approach, however, leads to control laws with increasing complexity as the time evolves. Another view that can be drawn from the deterministic theory, aiming at applications and control implementation, employs only the knowledge of the present observation to determine the control action. Those are the cases of decentralized and static output feedback systems [1], [2]. To describe these and other systems, particularly the ones where restriction exists on the observation of the system state x_k, we shall impose an *abstract* state-feedback structure for the controls u_k. To fix ideas, suppose that g belongs to a set of *actions* $\mathcal{G}$, and that K is a function that maps $\mathcal{G}$ to the space of real matrices of dimension $m \times n$. We then assume that u_k depends linearly on x_k as follows:

$$u_k = K(g)x_k, \quad \forall g \in \mathcal{G}, \quad \forall k \geq 0. \tag{2}$$

The variable g will be termed as *abstract control*, and the problem is to select an appropriate g such that the system (1) satisfies some desired characteristic. Observe that the abstract structure of (1)-(2) can be used to represent a broad range of linear control systems. For instance, in the particular case of decentralized control systems (see [1], [3]), the aim is to design a set of matrices $(G_1, \dots, G_n)$ such that

$$x_{k+1} = \left(A + \sum_{i=1}^{\eta} B_i G_i C_i \right) x_k + w_k, \; \forall k \geq 0, \; x(0) = x_0 \in \mathbb{R}^n, \tag{3}$$

evolves in such a manner that an index criterion is minimized. The correspondence between (1)-(2) and (3) follows easily by setting in (1)-(2), $B = I$, $g = (G_1, \dots, G_n)$, and $K(g) = \sum_{i=1}^{n} B_i G_i C_i$. Note in (3) that, if $\eta = 1$, then one retrieves the static-output feedback systems (see [2], [4]).

The main motivation of this paper is as follows. It is well-known that, if $\{w_k\}$ is a zero mean, gaussian process with covariance matrix equal to the identity, then the covariance matrix of x_k satisfies (see [5, Ch. 2], [6])

$$X_{k+1} = A_{g_k} X_k A'_{g_k} + \Sigma, \quad \forall g_k \in \mathcal{G}, \quad \forall k \geq 0, \tag{4}$$

where $A_g := A + BK(g)$ for all $g \in \mathcal{G}$, X_0 denotes the covariance matrix of x_0, and Σ is a nonnegative matrix. Let us assume in this discussion that X_k takes values in a set $\mathcal{X}$ only. Let $\mathcal{C} : \mathcal{X} \times \mathcal{G} \to \mathbb{R}_+$ be a suitable function for which the cost incurred at the k-stage is $\mathcal{C}(X_k, g_k)$. Then the long-run average cost is given by

$$J(\psi, X) := \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{C}(X_k, g_k), \tag{5}$$

with $\psi := \{g_0, g_1, \dots\}$, where $g_k \in \mathcal{G}$ for all $k \geq 0$, and $X_0 = X \in \mathcal{X}$. Let Ψ be the set of all feasible sequences ψ, and let $\Psi_s \subset \Psi$ be the set of all stationary sequences, so that if $\psi \in \Psi_s$ then $\psi = \{g, g, \dots\}$. The average cost

Research supported in part by Brazilian agencies FAPESP Grants 04/15672-5, 03/06736-7 and 04/06947-0; CNPq Grants 300721/86-2, 305657/2004-6 and 302345/2004-3.

Alessandro N. Vargas is with Universidade Tecnológica Federal do Paraná, UTFPR, Av. Alberto Carazzai 1640, 86300-000 Cornelio Procópio-PR, Brazil. E-mail: avargas@utfpr.edu.br

João B. R. do Val is with Universidade Estadual de Campinas, UNICAMP, Fac. de Engenharia Elétrica e de Computação, Depto. de Telemática, C.P. 6101, 13081-970 Campinas-SP, Brazil. E-mail: jbosco@dt.fee.unicamp.br

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

control problem is to find a sequence of abstract controls ψ^* such that

$$J(\psi^*,X) = \inf_{\psi\in\Psi} J(\psi,x) =: J^*(X), \quad \forall X \in \mathcal{X}. \tag{6}$$

It is clear that

$$\inf_{\psi\in\Psi} J(\psi,X) \leq \inf_{\psi\in\Psi_s} J(\psi,X), \quad \forall X \in \mathcal{X}. \tag{7}$$

The above inequality incites the following question.
(Q) *What are the conditions for which* (7) *holds with "=" in lieu of "≤"?*

There are two important reasons for willing the equality sign in (7). The first is motivated by the fact that stationary actions are preferred than non-stationary ones, mainly due to the easiness of implementation of the former when compared with the latter. The second reason is related to how to obtain a numerical solution for the long-run average cost control problem. For instance, to the best of the authors' knowledge, there is no method to solve numerically every control problem written as a static-output feedback one. Only particular cases of this problem can be dealt with by using numerical algorithms (see [2], [4], [7], [8] and the references therein). In view of this, we now discuss why the equality sign in (7) could be useful in the static-output feedback problem.

Suppose for the moment that one has a computational method for finding an optimal solution for the finite-horizon control problem

$$\inf_{\{g_0,\ldots,g_N\}} \sum_{k=0}^{N} \mathcal{C}(X_k,g_k). \tag{8}$$

Let $\{g_0^*,\ldots,g_N^*\}$ be an optimal solution for (8). Then one could hope that such time-dependent solution would approximate a stationary solution as long as the horizon N goes to infinity, or formally $\{g_0^*,g_1^*,\ldots,g_N^*\} \to \{g^*,g^*,\ldots\}$ as $N\to\infty$. Moreover, one expects that

$$\frac{1}{N}\sum_{k=0}^{N} \mathcal{C}(X_k^*,g_k^*) \to \frac{1}{N}\sum_{k=0}^{N} \mathcal{C}(X_k,g^*) \quad \text{ as } N\to\infty,$$

where $\{X_k^*\}$ is generated by $\{g_k^*\}$, and $\{X_k\}$ by $\{g^*\}$. Note, however, that the above approximation would be valid only if (7) holds with equality instead of strict inequality. Thus an important, and also an intriguing question, is how to assure the existence of optimal stationary solutions for infinite-horizon control problems. We shall restrict our analysis to the problem of minimizing the long-run average cost (6) subject to the dynamics (4).

The main contribution of this paper is to provide the conditions asked in (Q), with a minor conceptual modification. This occurs due to the following fact: the abstract control minimizing the right-hand side of (7) may depend on the state matrix sequence $\{X_k\}$, so that the infimum in the right-hand side of (7) is reached by a stationary function $g^*: \mathcal{X}\to\mathcal{G}$ instead of a stationary action $g^*\in\mathcal{G}$. Since the latter is a particular case of the former, we seek an optimal stationary function $g^*: \mathcal{X}\to\mathcal{G}$ to answer affirmatively (Q). We shall use results borrowed from the theory of Markov Control Processes (MCP), for which the question of existence of optimal stationary policies for the long-run average cost has been closely scrutinized (see the monographs [9], [10], and the articles [11], [12], [13], [14], [15] for further details). It is worthy to point out that the main technique used in the theory of MPC to prove the existence of optimal stationary policies for average cost problems is known as the *vanishing discount* approach, and the conditions we shall provide are based on it.

The paper is organized as follows. Section II is concerned with definitions, notations, and establishes the main results. In this section we introduce the notions of *policies*, the vanishing discount approach is revisited, and some important conditions are stated. We stress the assumptions of inf-compactness on the cost by stage, and the existence of optimal policies with exponential decay for the discounted problem. Finally, Section III presents some concluding remarks.

II. Preliminaries, notations, and main results

We denote respectively the real and natural numbers by $\mathbb{R}$ and $\mathbb{N}$. The normed linear space of all $n\times m$ real matrices is denoted by $\mathbb{R}^{n,m}$. The superscript $'$ indicates the transpose of a matrix. Let $\mathbb{S}^{n0}$ be the closed convex cone $\{U\in\mathbb{R}^{n,n} : U = U' \geq 0\}$; $\langle\cdot,\cdot\rangle$ will stand the inner product in $\mathbb{S}^{n0}$, and $\|\cdot\|$ will denote either the standard Euclidean norm in $\mathbb{R}^n$ or the Frobenius norm for matrices. We say that a matrix sequence $\{U_k; k\geq 0\}$ is bounded if $\sup_{k\in\mathbb{N}}\|U_k\| < \infty$.

The following definitions and conventions will apply throughout this paper.

(i) $\mathcal{X}$ and $\mathcal{G}$ are given sets referred to as *state space* and *abstract control space*, respectively. In particular, we assume $\mathcal{X}\subset\mathbb{S}^{n0}$.

(ii) For each $x\in\mathcal{X}$, there is given a nonempty measurable subset $\mathcal{G}(x)$ of $\mathcal{G}$. The set $\mathcal{G}(x)$ represents the set of *feasible abstract controls* or *actions* when the system is in state $x\in\mathcal{X}$, and with the property that the set

$$\mathbb{K} := \{(x,g) | x\in\mathcal{X}, g\in\mathcal{G}(x)\} \tag{9}$$

of feasible state-actions pairs is a measurable subset of $\mathcal{X}\times\mathcal{G}$.

(iii) We denote by $\mathcal{F}$ the set of all functions $f:\mathcal{X}\to\mathcal{G}$ such that $f(x)\in\mathcal{G}(x)$ for all $x\in\mathcal{X}$. The functions in $\mathcal{F}$ are called *selectors*. We denote by Π the set of all sequences $\pi = \{f_0,f_1,\ldots\}$ such that $f_k\in\mathcal{F}$ for all $k\geq 0$. Elements of Π are referred to *abstract policies*. Elements of Π of the form $\pi = \{f,f,\ldots\}$, where $f\in\mathcal{F}$, are referred to as *stationary abstract policies*.

(iv) Let $Q:\mathcal{G}\to\mathbb{S}^{n0}$ be a continuous, measurable function. The one-stage cost function $\mathcal{C}:\mathbb{K}\to\mathbb{R}_+$ is defined as follows:

$$\mathcal{C}(X,g) = \langle Q(g),X\rangle, \quad \forall (X,g)\in\mathbb{K}. \tag{10}$$

(v) Let $A : \mathcal{G} \to \mathbb{R}^{n,n}$ be a continuous, measurable function. In connection with (4), we define the following deterministic recurrence:

$$X_{k+1} = A(g_k)X_kA(g_k)' + \Sigma, \quad \forall k \geq 0, \quad X_0, \Sigma \in \mathcal{X}, \tag{11}$$

where the abstract control, applied at the k-th stage, is $g_k = f_k(X_k)$ whenever $\{f_k\} \in \Pi$.

We shall consider that the one-stage cost, defined in (10), satisfies in addition the following requirement.

Assumption 2.1: The one-stage cost $\mathcal{C} : \mathbb{K} \to \mathbb{R}_+$ is inf-compact on $\mathbb{K}$.

Recall that a function $v : \mathbb{K} \to \mathbb{R}$ is said to be inf-compact on $\mathbb{K}$ if, for each $x \in \mathcal{X}$ and $r \in \mathbb{R}_+$, the set $\{g \in \mathcal{G}(x) | v(x,g) \leq r\}$ is compact (see [9, p. 28], [16, p. 46]).

For sake of notational simplicity, the one-stage cost $\mathcal{C}(X_k, f_k(X_k))$ will be denoted by $\mathcal{C}(X_k, g_k)$ when $g_k = f_k(X_k)$ with $\{f_k\} \in \Pi$. This enables us to represent the long-run average cost by

$$J(\pi, X) := \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{C}(X_k, g_k), \tag{12}$$

when using the policy $\pi \in \Pi$ and initial state $X_0 = X \in \mathcal{X}$. The average cost control problem is to find a policy π^* such that

$$J(\pi^*, X) = \inf_{\pi \in \Pi} J(\pi, X) =: J^*(X), \quad \forall X \in \mathcal{X}. \tag{13}$$

The policy π^* satisfying (13) is referred to *average cost optimal*.

A. The vanishing discount approach

In this section we obtain conditions that assure the existence of a stationary optimal policy for the long-run average cost control problem. For this purpose we shall use results from the Markov Control Processes theory, in particular those related to the well-known vanishing discount approach [9], [10], [13], [17]. Let the discount criterion be defined as

$$V_\alpha(\pi, X) := \sum_{k=0}^{\infty} \alpha^k \mathcal{C}(X_k, g_k), \quad \pi \in \Pi, \quad X \in \mathcal{X}, \tag{14}$$

when using policy $\pi \in \Pi$, given the initial state $X_0 = X \in \mathcal{X}$, where α represents the *discount factor*. The α-discount abstract control problem is then given by

$$V_\alpha(\pi^*_\alpha, X) = \inf_{\pi \in \Pi} V_\alpha(\pi, X) =: V^*_\alpha(X), \quad \forall X \in \mathcal{X}, \tag{15}$$

and an abstract policy π^*_α satisfying (15) is said to be α-discount optimal.

Definition 2.1: Let $\Phi : \Pi \times \mathbb{N} \times \mathbb{N} \to \mathbb{M}^{n,n}$ be the following evolution operator:

$$\Phi(\pi, k, s) = \begin{cases} A(g_{k-1}) \cdots A(g_s) & \text{if } k > s \geq 0, \\ I & \text{otherwise,} \end{cases}$$

where $g_k = f_k(X_k)$ whenever $\pi = \{f_k\} \in \Pi$, and X_k is generated by (11).

With the above definition, we can introduce the following assumption.

Assumption 2.2: Let $\alpha_0 \in (0,1)$. Then, for all $\alpha \in [\alpha_0, 1)$, there exist $X_0 \in \mathcal{X}$ and a corresponding α-discount optimal abstract policy π^*_α such that

$$\|\Phi(\pi^*_\alpha, k+n, k)\| \leq M_\alpha \exp(-\xi_\alpha \cdot n), \quad \forall k > n \geq 0, \tag{16}$$

with $\sup_{\alpha_0 \leq \alpha < 1} M_\alpha < \infty$ and $\inf_{\alpha_0 \leq \alpha < 1} \xi_\alpha > 0$.

Assumption 2.2 roughly says that there exist α-discount optimal abstract policies π^*_α, with α within a neighborhood of 1, such that the evolution operator corresponding to π^*_α has a uniform exponential decay. Such assumption enable us to state the following main result.

Theorem 2.1: Suppose that Assumptions 2.1 and 2.2 hold. Then there exist a nonnegative constant ρ, a measurable function $h : \mathcal{X} \to \mathbb{R}_+$, and a selector $f^* \in \mathcal{F}$ such that

$$\begin{aligned} \rho + h(X) &\geq \min_{g \in \mathcal{G}(X)} \left[\mathcal{C}(X,g) + h\big(A(g)XA(g)' + \Sigma\big) \right] \\ &= \mathcal{C}(X, f^*(X)) + h\big(A(f^*(X))XA(f^*(X))' + \Sigma\big), \quad \forall X \in \mathcal{X}. \end{aligned} \tag{17}$$

Moreover, the following hold:

(i) The policy $f^\infty = \{f^*, f^*, \ldots\}$ is average cost optimal and ρ is the minimum average cost, that is,

$$J^*(X) = J(f^\infty, X) = \rho, \quad \forall X \in \mathcal{X}. \tag{18}$$

(ii) Any selector $f^* \in \mathcal{F}$ that satisfies (17) also satisfies the assertion in (i).

It is noteworthy from Theorem 2.1 that the average cost optimal value ρ in (18), and its corresponding stationary optimal policy f^∞, do not depend on the choice of the initial state $X_0 = X \in \mathcal{X}$.

B. Proof of Theorem 2.1

We shall show that the following two assertions hold:

(a) There exist a state $Z \in \mathcal{X}$ and numbers $\alpha_0 \in (0,1)$ and $L \geq 0$ such that

$$(1-\alpha)V^*_\alpha(Z) \leq L, \quad \forall \alpha \in [\alpha_0, 1). \tag{19}$$

(b) Set $h_\alpha(X) = V^*_\alpha(X) - V^*_\alpha(Z)$. Then there exists a measurable function $b : \mathcal{X} \to \mathbb{R}_+$ such that

$$0 \leq h_\alpha(X) \leq b(X), \quad \forall X \in \mathcal{X}, \quad \forall \alpha \in [\alpha_0, 1). \tag{20}$$

If one assumes that both (a) and (b) hold then the result of Theorem 2.1 follows straightforwardly from [9, Th. 5.4.3, p. 88] or [13, Th. 3.8]. Thus, it remain to show that (a) and (b) hold. For this, let us adopt the following convention. For any $\pi = \{f_k\} \in \Pi$, we assume that $g_k = f_k(X_k)$ with X_k generated by (11). Now, recalling the one-stage cost $\mathcal{C}(\cdot)$ in (10), we obtain the following preliminary result after some algebraic manipulation.

Lemma 2.1: There holds

$$\sum_{k=t}^{N-1} \alpha^k \mathcal{C}(X_k, g_k) = \langle L^{(\pi)}_{t,N}, X_t \rangle + \sum_{k=t+1}^{N-1} \alpha^k \langle L^{(\pi)}_{k,N}, \Sigma \rangle, \tag{21}$$

for all $N > 0$ and all $t = 0,\ldots,N$, where X_k satisfies (11) and

$$L^{(\pi)}_{k,N} = Q(g_k) + \alpha A(g_k)' L^{(\pi)}_{k+1,N} A(g_k), \quad k = 0,\ldots,N-1$$
$$L_{N,N} = 0. \tag{22}$$

The next result states that the sequence $\{L^{(\pi)}_{k,N}\}_{k=0}^N$ is bounded whenever π satisfies the exponential decay of Assumption 2.2.

Lemma 2.2: Suppose that Assumption 2.2 holds. Then, for all $\alpha \in [\alpha_0, 1)$, there exists a corresponding α-discount optimal policy π_α such that the following hold:

(i) The matrix sequence $\{P^\alpha_k\}$, with $P^\alpha_k := \lim_{N\to\infty} L^{(\pi_\alpha)}_{k,N}$, has a bound which does not depend on α.

(ii) There exists a function $\beta : (0,1) \to \mathbb{R}_+$ such that

$$\lim_{N\to\infty} \sum_{k=1}^{N-1} \alpha^k \langle L^{(\pi_\alpha)}_{k,N}, \Sigma\rangle = \beta(\alpha). \tag{23}$$

Proof: Assumption 2.2 implies that

$$\|\Phi(\pi_\alpha, k+n, k)\| \le M e^{-\xi n}, \quad k \ge n, \tag{24}$$

where $M := \sup_{\alpha_0 \le \alpha < 1} M_\alpha < \infty$ and $\xi := \inf_{\alpha_0 \le \alpha < 1} \xi_\alpha > 0$. Now, let $\{g^\alpha_k\}$ be the optimal sequence of actions corresponding to the optimal policy π_α, and let $X^{(\pi_\alpha)}_k$ be the recurrence (11) when it is evaluated to $\{g^\alpha_k\}$. It then follows from (24) that there exists a constant c_0 such that

$$\sup_{\alpha_0 \le \alpha < 1} \sup_{k\in\mathbb{N}} \|X^{(\pi_\alpha)}_k\| \le c_0. \tag{25}$$

We can conclude, from (25), that there exists a bounded set $\bar{\mathcal{X}} \subset \mathcal{X}$ so that $X^\alpha_k \in \bar{\mathcal{X}}$ for all $k \ge 0$ and all $\alpha \in [\alpha_0, 1)$. We now claim that there exists a constant c_1 such that

$$\sup_{\alpha_0 \le \alpha < 1} \sup_{k\in\mathbb{N}} \|Q(g^\alpha_k)\| \le c_1. \tag{26}$$

Indeed, since the inf-compact assumption on $\mathbb{K}$ (see Assumption 2.1) implies that the set

$$\lambda(r) := \{g \in \mathcal{G}(X) \mid \langle Q(g), X\rangle \le r, \forall X \in \bar{\mathcal{X}}\}$$

is compact for all $r \in \mathbb{R}_+$, then there exists a sufficiently large r_0 for which $g^\alpha_k \in \lambda(r_0)$ for all $k \ge 0$ and all $\alpha \in [\alpha_0, 1)$. This, together with the fact that $Q(\lambda(r))$ is a compact set (because Q is a continuous function) proves the claim.

Now, observe that (22) can be written equivalently to

$$L^{(\pi_\alpha)}_{k,N+1} = \sum_{j=k}^{N} \alpha^{N-j} Q(g_{N-j+k}) \Phi(\pi_\alpha, N-j+k, k). \tag{27}$$

Hence, we can see from (24), (26), and (27) that

$$\begin{aligned}
&\lim_{N\to\infty} \|L^{(\pi_\alpha)}_{k,N}\| \\
&= \lim_{N\to\infty} \left\| \sum_{j=k}^{N} \alpha^{N-j} Q(g_{N-j+k}) \Phi(\pi_\alpha, N-j+k, k) \right\| \\
&\le \lim_{N\to\infty} \sum_{j=k}^{N} \|Q(g_{N-j+k}) \Phi(\pi_\alpha, N-j+k, k)\| \\
&\le c_1 M \cdot \lim_{N\to\infty} \sum_{j=k}^{N} e^{-\xi(N-j)} \le c_1 M/(1-e^{-\xi}).
\end{aligned} \tag{28}$$

Thus $P^\alpha_k := \lim_{N\to\infty} L^{(\pi_\alpha)}_{k,N}$ is bounded for all $k \ge 0$ and all $\alpha \in [\alpha_0, 1)$, and this proves assertion *(i)*.

To prove assertion *(ii)*, notice that

$$\langle L^{(\pi_\alpha)}_{k,N}, \Sigma\rangle \le \|\Sigma\| \|L^{(\pi_\alpha)}_{k,N}\| \le \|\Sigma\| \lim_{N\to\infty} \|L^{(\pi_\alpha)}_{k,N}\| = \|\Sigma\| \|P^\alpha_k\|,$$

for all $k \ge 0$, and the conclusion follows using the result of assertion *(i)*. ■

Proof of Theorem 2.1 continued

The main argument in the proof of Theorem 2.1 now follows. Notice from (14), (15) and (21) that

$$V^*_\alpha(X) = \inf_{\pi\in\Pi} \left(\lim_{N\to\infty} \langle L^{(\pi)}_{0,N}, X\rangle + \sum_{k=1}^{N-1} \alpha^k \langle L^{(\pi)}_{k,N}, \Sigma\rangle \right). \tag{29}$$

Consider $X_0 \in \mathcal{X}$ as given by Assumption 2.2, and let $\pi_{\alpha,0} \in \Pi$ be an optimal abstract policy in (29) when $X = X_0$. We now show that there exists $M > 0$ such that

$$h_\alpha(X) = V^*_\alpha(X) - V^*_\alpha(X_0) \le M\|X\|, \quad \forall \alpha \in [\alpha_0, 1), \quad \forall X \in \mathcal{X}. \tag{30}$$

Indeed, by using (15) and (29) we have

$$\begin{aligned}
h_\alpha(X) = &\inf_{\pi\in\Pi} \left(\lim_{N\to\infty} \langle L^{(\pi)}_{0,N}, X\rangle + \sum_{k=1}^{N-1} \alpha^k \langle L^{(\pi)}_{k,N}, \Sigma\rangle \right) \\
&- V^*_\alpha(\pi_{\alpha,0}, X_0) \\
&\le \lim_{N\to\infty} \left[\langle L^{(\pi_{\alpha,0})}_{0,N}, X\rangle + \sum_{k=1}^{N-1} \alpha^k \langle L^{(\pi_{\alpha,0})}_{k,N}, \Sigma\rangle \right] \\
&- \lim_{N\to\infty} \left[\langle L^{(\pi_{\alpha,0})}_{0,N}, X_0\rangle + \sum_{k=1}^{N-1} \alpha^k \langle L^{(\pi_{\alpha,0})}_{k,N}, \Sigma\rangle \right],
\end{aligned} \tag{31}$$

where the inequality in (31) arises from an optimal argument. Observe now that, by Lemma 2.2(i), there exists $M > 0$ satisfying

$$\lim_{N\to\infty} \|L^{(\pi_{\alpha,0})}_{0,N}\| \le M, \quad \forall \alpha \in [\alpha_0, 1), \quad \forall k \ge 0. \tag{32}$$

Applying Lemma 2.2(ii) with (32) in the right-hand side of (31), we obtain

$$\begin{aligned}
h_\alpha(X) &\le \|X\| \cdot \lim_{N\to\infty} \|L^{(\pi_{\alpha,0})}_{0,N}\| + \beta(\alpha) \\
&\quad - \lim_{N\to\infty} \langle L^{(\pi_{\alpha,0})}_{0,N}, X_0\rangle - \beta(\alpha) \\
&\le M\|X\|,
\end{aligned}$$

which proves (30). Hence item (b) holds by letting $Z = X_0$ and $b(X) = M\|X\|$ for all $X \in \mathcal{X}$. We claim that (a) holds. Indeed, we can see from (29) and (32) that

$$\begin{aligned}(1-\alpha)V_\alpha^*(X_0) &= (1-\alpha)\lim_{N\to\infty}\left(\langle L_{0,N}^{(\pi_{\alpha,0})}, X_0\rangle + \sum_{k=1}^{N-1}\alpha^k \langle L_{k,N}^{(\pi_{\alpha,0})}, \Sigma\rangle\right)\\ &\le (1-\alpha)\sum_{k=0}^{\infty}\alpha^k\|M\|\cdot\max(\|X_0\|,\|\Sigma\|)\\ &= \|M\|\cdot\max(\|X_0\|,\|\Sigma\|),\end{aligned}$$

which shows the claim. This argument completes the proof of Theorem 2.1. □

C. Further conditions

In this section we provide some sufficient conditions for Assumption 2.2, motivated by the fact that such conditions, together with Assumption 2.1, will guarantee the existence of an optimal abstract stationary policy for the long-run average cost problem (6) (see Theorem 2.1).

To begin with, let us introduce the following notation. Let $\{A_k\}$ be any matrix sequence, and define

$$\Psi(k,s) = A_{k-1}A_{k-2}\cdots A_s, \quad \text{for each } k > s \ge 0, \qquad (33)$$

with $\Psi(s,s) = I$. In connection, we present below the controllability and observability concepts for time-varying systems [18], [19], [20], which will be useful in the proof of the next main theorem.

Definition 2.2: The pair (A_k, B_k) is uniformly controllable (or simply controllable) if there exists $T_c \ge 1$ and a real number $\sigma_c > 0$ such that, for all $k \ge T_c$,

$$\sum_{i=0}^{T_c-1}\Psi(k,k-i)B_{k-i-1}B_{k-i-1}'\Psi(k,k-i)' \ge \sigma_c I.$$

Definition 2.3: The pair (A_k, C_k) is uniformly observable (or simply observable) if there exists $T_o \ge 1$ and a real number $\sigma_o > 0$ such that, for all $k \ge 0$,

$$\sum_{i=0}^{T_o-1}\Psi(k+i,k)'C_{k+i}'C_{k+i}\Psi(k+i,k) \ge \sigma_o I.$$

Controllability and observability concepts allow us to relate a uniform bound on the cost $\mathcal{C}(\cdot)$ with a uniform exponential decay, as the one stated in (16). For this purpose, let us introduce the following assumption.

Assumption 2.3: Let $\alpha_0 \in (0,1)$. Then, for all $\alpha \in [\alpha_0, 1)$, there exist $X_0 \in \mathcal{X}$ and a corresponding α-discount optimal abstract policy $\pi_\alpha = \{f_k^\alpha\} \in \Pi$ such that

$$\mathcal{C}(X_k, f_k^\alpha(X_k)) \le \rho,$$

where ρ is some constant that does not depend on α, and X_k satisfies (11) with $A(f_k^\alpha(X_k))$ in place of $A(g_k)$. We are now ready to state the following theorem.

Theorem 2.2: Suppose that Assumptions 2.1 and 2.3 hold, and let $\{f_k^\alpha\} \in \Pi$ and X_k be as in Assumption 2.3. Define $A_k^\alpha = A(f_k^\alpha(X_k))$, and $Q_k^\alpha = Q(f_k^\alpha(X_k))$ for all $k \ge 0$ and all $\alpha \in [\alpha_0, 1)$, and let E be a matrix such that $\Sigma = EE'$. If the pair (A_k^α, E) is controllable, and $(A_k^\alpha, (Q_k^\alpha)^{\frac{1}{2}})$ is observable (with the numbers σ_c and σ_o in Definitions 2.2 and 2.3 do not depending on α), then the conclusions of Theorem 2.1 are valid.

Proof: Let $\alpha \in [\alpha_0, 1)$ be arbitrary but fixed. Let $\{Y_k\}$ be an output matrix sequence for such α, that is, let $Y_k := \Lambda_k X_k^\alpha \Lambda_k'$ where $\Lambda_k = (Q_k^\alpha)^{\frac{1}{2}}$ for all $k \ge 0$. Since $\|Y_k^{1/2}\| = \langle \Lambda_k, X_k\rangle \le \rho$, for all $k \ge 0$, then $\{\|Y_k^{1/2}\|\}$ is bounded. Thus, since $\{\|Y_k^{1/2}\|\}$ is bounded and the pair (A_k^α, Λ_k) is observable, we have that $\{\|X_k^{1/2}\|\}$ is bounded [21, Lemma 24]. The bounded sequence $\{\|X_k^{1/2}\|\}$, together with the controllability of the pair (A_k^α, E) yields [21, Lemma 25]

$$\|A_{k+n-1}^\alpha\cdots A_k^\alpha\| \le M_\alpha \exp(-\xi_\alpha \cdot n), \quad \forall k > n \in \mathbb{N},$$

where M_α and ξ_α are positive numbers. Using the fact that ρ, σ_c, and σ_o do not depend on α, one can shown straightforwardly that there exist $M, \xi > 0$ such that

$$M_\alpha \le M < \infty, \quad 0 < \xi \le \xi_\alpha, \quad \forall \alpha \in [\alpha_0, 1).$$

Hence Assumption 2.2 holds, and the result then follows from Theorem 2.1. ■

III. Concluding remarks

This paper has shown that there exists an stationary optimal policy for the long-run average cost control problem provided that Assumptions 2.1 and 2.2 are satisfied. The derived results consider a discrete-time linear system under an abstract state-feedback structure (see (2) and (9)). Our results rely on (i) the vanishing discount approach, and (ii) the existence of an exponential decay for the control action that minimizes the α-discount problem, with α within a neighborhood of 1. From these assumptions, Theorem 2.1 shows that an stationary optimal policy $f^\infty = \{f^*, f^*, \ldots\}$ for the average cost problem exists. Moreover, Theorem 2.2 shows that the results of Theorem 2.1 apply, under the hypothesis of controllability, observability, together with an uniform bound on the cost $\mathcal{C}(\cdot)$, when it is evaluated to α-discount optimal policies π_α.

References

[1] N. R. Sandell, M. Athans, P. Varaiya, and M. G. Safonov, "Survey of decentralized control methods for large scale systems," *IEEE Trans. Automat. Control*, vol. 23, no. 2, pp. 108–128, 1978.

[2] V. Syrmos, C. Abdallah, P. Dorato, and K. Grigoriadis, "Static output feedback—a survey," *Automatica*, vol. 33, pp. 125–137, 1997.

[3] B. D. O. Anderson and J. B. Moore, "Time-varying feedback laws for decentralized control," *IEEE Trans. Automatic Control*, vol. 26, no. 5, pp. 1133–1138, 1981.

[4] J. C. Geromel, C. C. de Souza, and R. E. Skelton, "Static output feedback controllers: stability and convexity," *IEEE Trans. Automatic Control*, vol. 43, no. 1, pp. 120–125, 1998.

[5] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Prentice-Hall, Englewood Cliffs, N.J., 1979.

[6] M. H. A. Davis and R. B. Vinter, *Stochastic Modelling and Control.* Chapman and Hall, London, 1985.

[7] V. Blondel and J. N. Tsitsiklis, "NP-hardness of some linear control design problems," *SIAM J. Control Optim.*, vol. 35, no. 6, pp. 2118–2127, 1997.

[8] D. Henrion and J. B. Lasserre, "Convergent relaxations of polynomial matrix inequalities and static output feedback," *IEEE Trans. Automatic Control*, vol. 51, no. 2, pp. 192–202, 2006.
[9] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria.* Springer-Verlag, New York, 1996.
[10] ——, *Further Topics in Discrete Time Markov Control Processes.* Springer-Verlag, New York, 1999.
[11] O. Hernández-Lerma and R. Romera, "Limiting discounted-cost control of partially observable stochastic systems," *SIAM J. Control Optim.*, vol. 40, no. 2, pp. 348–369, 2001.
[12] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, "Discrete-time controlled Markov processes with average cost criterion: A survey," *SIAM J. Control Optim.*, vol. 31, no. 2, pp. 282–344, 1993.
[13] M. Schal, "Average optimality in dynamic programming with general state space," *Math. of Operations Res.*, vol. 18, pp. 163–172, 1993.
[14] S. P. Meyn, "The policy iteration algorithm for average reward Markov decision processes with general state space," *IEEE Trans. Automat. Control*, vol. 42, no. 12, pp. 1663–1680, 1997.
[15] E. A. Feinberg and M. E. Lewis, "Optimality inequalities for average cost Markov decision processes and the stochastic cash balance problem," *Math. Oper. Res.*, vol. 32, no. 4, pp. 769–783, 2007.
[16] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete Time Case.* Academic Press, 1978.
[17] O. Hernández-Lerma and J. B. Lasserre, "Average cost optimal policies for Markov control processes with Borel state space and unbounded costs," *Systems Control Lett.*, vol. 15, pp. 349–356, 1990.
[18] R. E. Kalman, "Contribution to the theory of optimal control," *Bol. Soc. Mat. Mexicana*, vol. 5, no. 2, pp. 102–119, 1960.
[19] L. M. Silverman and B. D. O. Anderson, "Controllability, observability and stability of linear systems," *SIAM J. Control Optim.*, vol. 6, no. 1, pp. 121–130, 1968.
[20] B. D. O. Anderson and J. B. Moore, "New results in linear system stability," *SIAM J. Control*, vol. 7, no. 3, pp. 398–414, 1969.
[21] A. N. Vargas and J. B. R. do Val, "Average cost and stability of time-varying linear systems," *Preprint*