

Stability Estimation for Markov Control Processes with Discounted Cost

Jaime Eduardo Martínez-Sánchez

Departamento de Mercadotecnia y Análisis de Datos, Tecnológico de Monterrey, Campus Estado de México (ITESM-CEM), Atizapán de Zaragoza, Estado de México, Mexico

Email: meduardo@tec.mx

How to cite this paper: Martínez-Sánchez, J.E. (2020) Stability Estimation for Markov Control Processes with Discounted Cost. *Applied Mathematics*, 11, 491-509. <https://doi.org/10.4236/am.2020.116036>

Received: May 18, 2020

Accepted: June 19, 2020

Published: June 22, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This article explores controllable Borel spaces, stationary, homogeneous Markov processes, discrete time with infinite horizon, with bounded cost functions and using the expected total discounted cost criterion. The problem of the estimation of stability for this type of process is set. The central objective is to obtain a bounded stability index expressed in terms of the Lévy-Prokhorov metric; likewise, sufficient conditions are provided for the existence of such inequalities.

Keywords

Discrete-Time Markov Control Process, Expected Total Discounted Cost, Stability Index, Probabilistic Metric, Lévy-Prokhorov Metric

1. Introduction

Let on a Borel space (X, \mathbb{B}_X) two following Markov control processes be given:

$$x_t = F(x_{t-1}, a_t, \xi_t), \quad t = 1, 2, \dots \quad (1.1)$$

$$\tilde{x}_t = F(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t), \quad t = 1, 2, \dots \quad (1.2)$$

where $a_t \in A(x_{t-1}) \subset A$, $\tilde{a}_t \in A(\tilde{x}_{t-1}) \subset A$ are the controls forming the control policies $\pi = (a_1, a_2, \dots)$, $\tilde{\pi} = (\tilde{a}_1, \tilde{a}_2, \dots)$ (see [1] [2] for definitions); $\{\xi_t\}, \{\tilde{\xi}_t\}$ are sequences of independent and identically distributed (i.i.d.) random vectors in a separable metric space (S, r) . In what follows the distributions of $\xi_1, \tilde{\xi}_1$ are denoted by D_ξ and $D_{\tilde{\xi}}$ respectively. Let c be a given bounded measurable one-step cost function; for any initial state $x \in X$ and control policy $\pi \in \Pi$ (Π is the set of all control policies see [1]), the expected total α -discounted cost

criterion areas follows:

$$V_\alpha(x, \pi) := \mathbb{E}_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(x_{t-1}, a_t), \quad (1.3)$$

$$\tilde{V}_\alpha(x, \pi) := \mathbb{E}_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(\tilde{x}_{t-1}, \tilde{a}_t), \quad (1.4)$$

Under assumptions 3.1 and 3.2 given in section 3, there exist stationary optimal policies f_* and \tilde{f}_* such that

$$\begin{aligned} V_\alpha^* &= V_\alpha(f_*) = \inf_{\pi \in \Pi} V_\alpha(x, \pi) \\ \tilde{V}_\alpha^* &= \tilde{V}_\alpha(\tilde{f}_*) = \inf_{\pi \in \Pi} \tilde{V}_\alpha(x, \pi) \end{aligned} \quad (1.5)$$

To set the stability estimation problem, first, suppose that process given in Equation (1.2) is interpreted as an “available approximation” to process given in Equation (1.1), *i.e.*, $D_{\tilde{\xi}}$ is an approximation to D_{ξ} .

Second, the policy \tilde{f}_* (optimal with respect to Equation (1.4)) is applied to control the “original process” given in Equation (1.1) (instead of “unavailable” optimal policy f_*).

Following the definition given in [3] [4] [5] [6] [7], we introduce the stability index:

$$\Delta_{V_\alpha} := V_\alpha(\tilde{f}_*) - V_\alpha(f_*) \geq 0,$$

where $V_\alpha(\cdot)$ is the value function defined in Equation (1.5). This definition means that Δ_{V_α} represents an extra cost paid for using \tilde{f}_* instead of the optimal policy f_* .

Under certain Lipschitz conditions it was proved (for the processes with bounded costs c) that

$$\Delta_{V_\alpha} = \bar{K} \ell \pi(D_\xi, D_{\tilde{\xi}}) \quad (1.6)$$

where \bar{K} is an explicitly calculated constant, and $\ell \pi$ is the Lévy-Prokhorov metric (see Section 2 for definition). The convergence in $\ell \pi$ is equivalent to the weak convergence plus the convergence of first absolute moments (see [8]).

Inequalities as given in Equation (1.6) have been developed with other types of metrics (Kantorovich, total variation, etc.) and optimization criteria (the average cost) see e.g. in [5] [7] [9] [10]. Other types of criteria used to obtain the stability of the process can be consulted in [11] [12] [13].

The aim of the present paper is making advantage of boundedness of c and using the well-known contractive properties of the operators related to the expected total discount cost optimality equations to prove the “stability inequality” as in Equation (1.6) with the Lévy-Prokhorov distance on its right-hand side.

This paper is organized as follows: Section 2 defines the control Markov model and the problem of its stability. Section 3 presents the Lipschitz conditions and the assumptions to guarantee the existence of an optimal control to the Markov control process as well as the main result of this work, the Theorem 3.1, which establishes the conditions to achieve the stability. Section 4 is presented a

couple of application examples, for which the assumptions are validated and then the result obtained in the Theorem 3.1 is applied. Finally, section 5 has presented the proof of the Theorem 3.1 as well as a couple of lemmas that are required for this proof.

2. Setting of the Problem

In standard way (see for instance [1] [14]), it will denote a Markov Control Process (MCP) indiscrete time with infinite horizon, stationary and homogeneous as the following fivefold:

$$M = (X, A, \{A(x) : x \in X\}, p, c) \quad (2.1)$$

where will be assumed that the controllable process components M have the following characteristics:

- The space state X is a metric space with a metric ρ and \mathbb{B}_X denotes the sigma-algebra;
- The actions space A is a metric space with a metric l ;
- The set of admissible actions $A(x)$ is compact for every $x \in X$;
- The pairs set of admissible state-actions $\mathbb{K} = \{(x, a) \in X \times A : a \in A, x \in X\}$ is anon-empty (and measurable) Borel subset of the set $X \times A$ and it is equipped with the metric $\nu = \max\{\rho, l\}$;
- p is a stochastic kernel in X given \mathbb{K} . This stochastic kernel specifies the transition probability:

$$p = \mathbb{P}(B|x, a) \quad (2.2)$$

where $B \in \mathbb{B}_X$ and $(x, a) \in \mathbb{K}$.

- Finally, $c : \mathbb{K} \rightarrow \mathbb{R}$ is a bounded and measurable function called a step cost function.

On the other hand, in many applications the evolution of the MCP given in Equation (2.1) is specified by the following model:

$$x_t = F(x_{t-1}, a_t, \xi_t), \quad t = 1, 2, \dots \quad (2.3)$$

where x_0 represents the initial state and $\{\xi_t\}$ it is a sequence of i.i.d. random vectors that take values in any Borel space S with a common distribution D_ξ . In fact, it is considered that S is a metric space equipped with a metric r and $F : \mathbb{K} \times S \rightarrow X$ is a measurable function. The expression given in Equation (2.3) will be denoted as the original process.

Let $x \in X$ be the initial state and $\pi \in \Pi$ the applied control policy, (Π is the set of all control policies, see [1] [14] for definitions), then the performance criterion called expected total α -discounted cost is defined as usual, by the following functional:

$$V_\alpha(x, \pi) := \mathbb{E}_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(x_{t-1}, a_t) \quad (2.4)$$

where $\alpha \in (0, 1)$ is a fixed discount coefficient; \mathbb{E}_x^π denotes the expected value corresponds to the distribution of the process $\{x_t\}$ with the initial state $x \in X$

and the control policy $\pi \in \Pi$ applied.

Now, the function $V_\alpha^*(x) := \inf_{\pi \in \Pi} V_\alpha(x, \pi)$ with $x \in X$ is called the value function and a control policy π_* (provided it exists) is called optimum (with respect to the criterion V_α) if it meets the following:

$$V_\alpha^*(x) = V_\alpha(x, \pi_*) = \inf_{\pi \in \Pi} V_\alpha(x, \pi), \quad x \in X \quad (2.5)$$

Later, conditions will be imposed that will guarantee the existence of an optimal stationary policy $\pi_* = f_* = (f, f, \dots, f, \dots)$ for Equation (2.5), (see [14]).

The stability index and its estimation problem. Estimation of the stability problem arises when there is uncertainty about the likelihood of transition p defined in Equation (2.2). The original task of controller consists of the search (or approach) of the optimal policy π_* that satisfies Equation (2.5) for the original process. In many applications, this task cannot be accomplished directly because, among other reasons, any of the following:

1) Frequently p or some of its parameters are unknown to the controller and this transition probability is estimated using some statistical procedures (from observations). With the results of these estimates another transition probability is generated \tilde{p} , that is interpreted as an approximation accessible to the unknown p .

2) There are situations in which p is known but too complicated to have a hope to solve the problem of optimization of control policy. In such cases, sometimes p is replaced by “a theoretical approach” \tilde{p} , resulting in a controllable process with a more simple structure.

In both cases, in the optimization policies the controller is to work with the controllable Markov process $\tilde{M} = (X, A, \{A(x) : x \in X\}, \tilde{p}, c)$ defined by the accessible transition probability \tilde{p} . This means that instead of the original process x_t , given in Equation (2.3), the controller uses an approximate process given by the following equation:

$$\tilde{x}_t = F(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t), \quad t = 1, 2, \dots, \text{ with } x_0 \in X \text{ given} \quad (2.6)$$

where $\tilde{x}_t, \tilde{x}_{t-1} \in X$ are states of the processes; $\tilde{a}_t \in A(\tilde{x}_{t-1})$ is an action of the corresponding state; and $\{\tilde{\xi}_t\}$ is a sequence of random vector i.i.d. with values in S . The only difference between the given processes in Equations (2.3) and (2.6) is possible, the different distributions D_ξ and $D_{\tilde{\xi}}$ from the random vector $\{\xi_t\}$ and $\{\tilde{\xi}_t\}$ respectively.

Changing x_t for \tilde{x}_t in Equations (2.4) and (2.5), it is defined as the corresponding optimization criterion \tilde{V}_α for approximate processing \tilde{M}

$$\tilde{V}_\alpha(x, \pi) := \mathbb{E}_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(\tilde{x}_{t-1}, \tilde{a}_t), \text{ with } x_0 \in X.$$

Suppose now, that it is possible (at least theoretically) to find an optimal policy $\tilde{\pi}_*$ for the process \tilde{M} , *i.e.*, the value for the approximate process function is defined as

$$\tilde{V}_\alpha^*(x, \tilde{\pi}_*) = \tilde{V}_\alpha^*(x) := \inf_{\pi \in \Pi} \tilde{V}_\alpha(x, \pi), \quad x \in X \quad (2.7)$$

The control policy $\tilde{\pi}_*$ in Equation (2.7) is used as approximation to the nonaccessible optimal policy π_* (assuming it exists). In other words, the policy $\tilde{\pi}_*$ is used to control the original process M instead of the unknown policy π_* .

The increase in the cost for such an approach is estimated by means of the following stability index, (see [3] [4]):

$$\Delta_{V_\alpha}(x) := V_\alpha(x, \tilde{\pi}_*) - V_\alpha^*(x) \equiv V_\alpha(x, \tilde{\pi}_*) - V_\alpha(x, \pi_*) \geq 0, \quad x \in X \quad (2.8)$$

As proposed in [5] [6], the estimation of the stability problem consists of the search of some inequalities of the following type (stability inequalities):

$$\Delta_{V_\alpha}(x) \leq \bar{K}(x) \psi[\mu(p, \tilde{p})], \quad x \in X \quad (2.9)$$

where:

$\mu(p, \tilde{p})$ is a “distance” between the probabilities of transition p and \tilde{p} (expressed in terms of a probabilistic metric).

$\psi[\mu(p, \tilde{p})]$ is a continuous function such that $\psi(s) \rightarrow 0$ when $s \rightarrow 0$; and $\bar{K}(x), x \in X$ is a function with values calculated explicitly.

The results presented in [4] [5] provide inequalities such as the one given in the inequality (2.9) using $\psi(s) = s^\gamma$ for $0 < \gamma \leq 1$, and the so-called “strong metrics”: total variation metric and the weighted total variation metric.

The aim of this article is obtaining inequalities of stability such as given in the inequality (2.9) with $\psi(s) = s$ and the use of “metric weak” probabilistic, specifically, the Lévy-Prokhorov metric ($\ell\pi$).

For instance, the Theorem 3.1 presented in the next section, see inequalities (3.1) and (3.2), ensure that under appropriate conditions it complies

$$\Delta_{V_\alpha}(x) \leq \bar{K}(x) \ell\pi(D_\xi, D_{\tilde{\xi}}), \quad x \in X \quad (2.10)$$

where:

$$\ell\pi(D_\xi, D_{\tilde{\xi}}) := \inf \left\{ \varepsilon > 0 : D_\xi(A) \leq D_{\tilde{\xi}}(A^\varepsilon) + \varepsilon, \right. \\ \left. D_{\tilde{\xi}}(A) \leq D_\xi(A^\varepsilon) + \varepsilon \text{ for every } A \in \mathbb{B}_S \right\}$$

is the Lévy-Prokhorov metric; $A^\varepsilon := \{s \in S : r(s, A) < \varepsilon\}$ and \mathbb{B}_S denotes the sigma-algebra of Borel of metric space (S, r) .

It is well-known (see [15]) that $\ell\pi$ metrizes weak convergence in any separable metric space. A succession of random vectors that converge under the metric $\ell\pi$, converges weakly.

3. Assumptions and Results

The Hausdorff distance (h) between compact subsets B, C of the metric space (A, l) is given by

$$h(B, C) := \max \left\{ \sup_{x \in B} l(x, C), \sup_{y \in C} l(B, y) \right\},$$

where $l(x, C) = \inf_{z \in C} l(x, z)$.

Likewise the so-called “strong metric”, the total variation metric ($\mathbb{V}(\cdot, \cdot)$) is

given by

$$\mathbb{V}(D_{\xi}, D_{\tilde{\xi}}) := \sup \left\{ \left| \mathbb{E} \left[\phi(\xi) - \phi(\tilde{\xi}) \right] \right| : \|\phi\|_{\infty} < 1 \right\},$$

where $D_{\xi}, D_{\tilde{\xi}}$ are in the space of probability distributions over (S, \mathbb{B}_S) and $\|\cdot\|_{\infty}$ is the supremum norm. Of course under $S \equiv \mathbb{R}$, then

$$\mathbb{V}(D_{\xi}, D_{\tilde{\xi}}) = \int_{\mathbb{R}} |D_{\xi}(t) - D_{\tilde{\xi}}(t)| dt.$$

On the other hand, one of the metrics called “weak” is the Kantorovich metric (κ)

$$\kappa(D_{\xi}, D_{\tilde{\xi}}) := \sup_{\phi \in \mathbb{L}} \left\{ \mathbb{E} \left| \phi(\xi) - \phi(\tilde{\xi}) \right| \right\}$$

where the function ϕ it is of Lipchitz, namely, the set \mathbb{L} is defined as

$$\mathbb{L} := \left\{ \phi : |\phi(s) - \phi(s')| \leq r(s, s'), s, s' \in S \right\}.$$

It is well known (see [9]) that in the case of $S \equiv \mathbb{R}^m$, it is true that $\kappa(D_{\xi_n}, D_{\xi}) \rightarrow 0$ if and only if $\xi_n \rightarrow \xi$ (weak convergence) and that $\mathbb{E}|\xi_n| \Rightarrow \mathbb{E}|\xi|$.

In the remainder of the article, it will be denoted by \mathbf{B} to the Banach space of all measurable functions $u : X \rightarrow \mathbb{R}$ for which the norm $\|u\|_{\infty} = \sup_{x \in X} \{ |u(x)| \}$ is finite.

The first set of technical assumptions is required to ensure the existence of minimizers in the value functions of the original and the approximate model, see [16].

Assumption 3.1.

1) The set A is compact for each $x \in X$; also the mapping of values set as $x \in A(x)$ is upper semicontinuous with respect to the Hausdorff metric.

2) The one-step cost function $c : \mathbb{K} \rightarrow \mathbb{R}$ is bounded, namely $|c(k)| \leq b$ for each $k \in \mathbb{K}$, $b \in \mathbb{R}$; and for each $x \in X$, the one-step cost function $c(k) = c(x, a)$ is lower semicontinuous in A .

3) For each continuous function bounded $u : X \rightarrow \mathbb{R}$, the functions

$$u_1(x, a) := \mathbb{E}u[F(x, a, s)];$$

$$u_2(x, a) := \mathbb{E}u[F(x, a, \tilde{s})],$$

with $(x, a) \in \mathbb{K}$, $s, \tilde{s} \in S$ are continuous in \mathbb{K} .

The second set of assumptions imposes the “Lipschitz conditions” on the one-step cost function as well as on the transition probabilities of the original and approximate processes.

Assumption 3.2.

There are finite constants b, L_0, L_1, L_2, L_3 such that the following is true:

- 1) $|c(k)| \leq b < \infty$ for each $k \in \mathbb{K}$;
- 2) $|c(k) - c(k')| \leq L_0 v(k, k')$ for all $k, k' \in \mathbb{K}$;
- 3) $h(A(x), A(x')) \leq L_1 \rho(x, x')$ for all $x, x' \in X$ where $h(\cdot, \cdot)$ is the Hausdorff metric;

- 4) $\mathbb{V}(F(k, s), F(k', s)) \leq L_2 v(k, k')$ for all $x, x' \in X$, $s \in S$ where $\mathbb{V}(\cdot, \cdot)$ is the total variation metric;
- 5) $\rho(F(k, s), F(k, s')) \leq L_3 r(s, s')$ for all $k \in \mathbb{K}$, $s, s' \in S$;
- 6) For each $x \in X$, $s \in S$ and the bounded function $u: X \rightarrow \mathbb{R}$, then the function $a \rightarrow \mathbb{E}u[F(x, a, s)]$ is lower semicontinuous in $A(x)$.

For a proof of the following proposition, see [16].

Proposition 1 (Well-known result). Under the assumptions 3.1 and 3.2, for the control processes given in Equations (2.3) and (2.6) there are optimal stationary control policies denoted by $f_* = \{f_*, f_*, \dots\}$ and $\tilde{f}_* = \{\tilde{f}_*, \tilde{f}_*, \dots\}$ respectively, such that $V_\alpha(x, f_*)$ and $\tilde{V}_\alpha(x, \tilde{f}_*)$ do not depend on the initial state $x \in X$ and

$$V_\alpha^* = V_\alpha(x, f_*) = \inf_{\pi \in \Pi} V_\alpha(x, \pi), \quad x \in X;$$

$$\tilde{V}_\alpha^* = \tilde{V}_\alpha(x, \tilde{f}_*) = \inf_{\pi \in \Pi} \tilde{V}_\alpha(x, \pi), \quad x \in X.$$

In addition, the corresponding value functions $V_\alpha^*, \tilde{V}_\alpha^* \in \mathbf{B}$. In particular, for each fixed $(x, a) \in \mathbb{K}$, expected values $\mathbb{E}V_\alpha^*[F(x, a, s)]$ and $\mathbb{E}\tilde{V}_\alpha^*[F(x, a, s)]$ are well defined.

Now, we are in position to formulate the main result of the paper.

Theorem 3.1. Under the assumptions 3.1 and 3.2, the stability index given in Equation (2.8) meets the following inequality:

$$\sup_{x \in X} \Delta_{V_\alpha}(x) \leq \bar{K}_\alpha \ell \pi(D_\xi, D_{\tilde{\xi}}), \quad (3.1)$$

where the stability constant is

$$K_\alpha = \frac{4\alpha}{(1-\alpha)^3} [b + L_3(1+L_1)(L_0(1-\alpha) + \alpha b L_2)] \quad (3.2)$$

Note that if $\alpha \rightarrow 1$, then the constant \bar{K}_α in the inequality (3.2) it is of order $O(1-\alpha)^3$.

4. Some Examples

4.1. The Process of Regularization of the Water Level in a Dam

An important application of control problems (deterministic and stochastic) are those related to water reserve operations. An excellent introduction to many of these problems, including the connection between these and inventory systems, is given in [17].

In the simplest case of regularization of the water level in a dam, the following modeling can be used for the original process:

$$x_t = \min\{x_{t-1} + \xi_t - a_t; U\}, \quad t = 1, 2, \dots \quad (4.1)$$

and the respective approximate model remains as

$$\tilde{x}_t = \min\{\tilde{x}_{t-1} + \tilde{\xi}_t - \tilde{a}_t; U\}, \quad t = 1, 2, \dots \quad (4.2)$$

In this model, the state variable x_t represents the level of the stock (volume) of water that the dam has at the beginning of the period t ; the control a_t is the

amount of water that is released from the dam for family consumption, irrigation, electric power, etc. during the period t , and the “disturbance” ξ_t is the amount of water that the dam receives, randomly, viarain for instance.

In this example, we get $X = [0, U]$, $S = [0, \infty)$, $A(x) = [0, x]$, with $x \in X$, where U is the maximum capacity of the dam.

Let $0 \leq c(x, a) \leq b < \infty$ be the cost paid for the released water service, for example, can be made use of a cost function given by $c(x, a) = c_0 a$, proportional to water consumption and where c_0 would represent the cost of a unit of water.

To ensure compliance with assumption 3.2 for this example, it is admitted that the following conditions are met:

- C1. The cost for one step $c(x, a)$ satisfies the assumption 3.2 clauses (1) and (2).
- C2. The random variable ξ has a density g_ξ , which is:
 - 1) Bounds by a constant $M_g < \infty$;
 - 2) Satisfies the condition of Lipschitz with a constant L_g .

For $A(x) = [0, x]$, the clause (3) in the assumption 3.2 is verified directly (using the Hausdorff metric definition) with the constant $L_1 = 1$. Now, denoting by $y := x - a$, it is easy to see that for each y fixed, the function $F(x, a, s) = \min\{y + s, U\}$ is Lipschitz in S with the constant 1. Then the clause (5) of this assumption is complied with $L_3 = 1$. Next, the clauses will be verified (4) of assumption 3.2.

Denoting by $y := x - a$ and $y' := x' - a'$ with $x, x' \in [0, U]$, $a \in [0, x]$, $a' \in [0, x']$, consider the following random variables:

$$\zeta(y) := \min\{y + \xi, U\},$$

$$\zeta(y') := \min\{y' + \xi, U\}.$$

Since

$$|y - y'| \leq |x - x'| + |a - a'| \leq 2v(k, k'),$$

it is enough to prove that for a constant \tilde{L} the following inequality is met

$$\mathbb{V}(\zeta(y), \zeta(y')) \leq \tilde{L}|y - y'|, \quad (4.3)$$

At the time you will see that, according to the definition of the total variation metric, to prove the inequality (4.3) it must be proved that for each measurable function $\phi: S \rightarrow \mathbb{R}$, with $\|\phi\|_\infty \leq 1$ it is true that

$$|\mathbb{E}\phi[\zeta(y)] - \mathbb{E}\phi[\zeta(y')]| \leq \tilde{L}|y - y'|.$$

Now then

$$\mathbb{E}\phi[\zeta(y)] = \mathbb{E}\{\phi[\zeta(y)]; y + \xi < U\} + \mathbb{E}\{\phi[\zeta(y)]; y + \xi \geq U\},$$

where for a random variable η , we get that $\mathbb{E}\{\eta; A\} = \mathbb{E}\eta I_A$.

Using the same representation for $\mathbb{E}\phi[\zeta(y')]$, we get that

$$\begin{aligned} & \left| \mathbb{E}\phi[\zeta(y)] - \mathbb{E}\phi[\zeta(y')] \right| \\ & \leq \left| \mathbb{E}\{\phi[\zeta(y)]; y + \xi < U\} - \mathbb{E}\{\phi[\zeta(y)]; y' + \xi < U\} \right| \\ & \quad + \left| \mathbb{E}\{\phi[\zeta(y)]; y + \xi > U\} - \mathbb{E}\{\phi[\zeta(y)]; y' + \xi > U\} \right|, \end{aligned}$$

then

$$\left| \mathbb{E}\phi[\zeta(y)] - \mathbb{E}\phi[\zeta(y')] \right| \leq I_1(y, y') + I_2(y, y') \quad (4.4)$$

For the second term on the right side of the last inequality, we get that

$$\begin{aligned} I_2(y, y') &= \left| \phi(U) \mathbb{E}(\xi > U - y) - \phi(U) \mathbb{E}(\xi > U - y') \right| \\ &= \left| \phi(U) \left| \int_{U-y}^{\infty} g_{\xi}(s) ds - \int_{U-y'}^{\infty} g_{\xi}(s) ds \right| \right|. \end{aligned} \quad (4.5)$$

Let for instance be $y > y'$. Then from Equation (4.5) and the condition C2, we get that

$$I_2(y, y') = \int_{U-y}^{U-y'} g_{\xi}(s) ds \leq M_g |y - y'|. \quad (4.6)$$

Let $z \in (0, U)$ be an arbitrary but fixed number and dz that denotes an infinitesimal interval with center in z . Since

$$\{y + \xi \in dz\} \subset \{y + \xi < U\},$$

then

$$\mathbb{P}(\xi \in dz - y, y + \xi < U) = \mathbb{P}(\xi \in dz - y) = g_{\xi}(z - y) dz.$$

Similarly

$$\mathbb{P}(\xi \in dz - y', y' + \xi < U) = g_{\xi}(z - y') dz.$$

Then in the inequality (4.4) (taking into account that $g_{\xi}(x) = 0$ for $x < 0$):

$$I_1(y, y') = \left| \int_y^U \phi(z) g_{\xi}(z - y) dz - \int_{y'}^U \phi(z) g_{\xi}(z - y') dz \right|,$$

or then, assuming for example that $y > y'$, we get that

$$\begin{aligned} I_1(y, y') &= \left| \int_y^U \phi(z) g_{\xi}(z - y) dz - \int_{y'}^y \phi(z) g_{\xi}(z - y') dz - \int_y^U \phi(z) g_{\xi}(z - y') dz \right| \\ &\leq \left| \int_{y'}^y \phi(z) g_{\xi}(z - y') dz \right| + \int_y^U |\phi(z)| |g_{\xi}(z - y) - g_{\xi}(z - y')| dz \\ &\leq M_g |y - y'| + UL_g |y - y'| \end{aligned} \quad (4.7)$$

(Applying the conditions C1 and C2).

Joining inequalities (4.4), (4.6) and (4.7) is obtained the inequality (4.3) with $\tilde{L} = 2M_g + UL_g$.

Finally it has been established that for this example the clause (4) of assumption 3.2 is met with $L_2 = 2(2M_g + UL_g)$. Following similar arguments can be shown that the clause (6) of assumption 3.2 is also true. Therefore, in this example inequality (3.1) can be applied to the Theorem 3.1, obtaining the following:

$$\sup_{x \in [0, U]} \Delta_{V_\alpha}(x) \leq \bar{K}_\alpha \ell \pi(D_\xi, D_{\tilde{\xi}}) \quad (4.8)$$

where

$$\bar{K}_\alpha = \frac{4\alpha}{(1-\alpha)^3} \left[b + 2L_0(1-\alpha) + 4\alpha b(2M_g + UL_g) \right] \quad (4.9)$$

On the other hand, the distance $\ell \pi(D_\xi, D_{\tilde{\xi}})$ given in the inequality (4.8) is very difficult to calculate. Therefore, the result given in the inequality (4.8) can be expressed in terms of other probabilistic metrics as it's shown in the following:

❖ **Total variation metric.** Using the well-known relationship $\ell \pi < \mathbb{V}$, see [18], between the metrics of Lévy-Prokhorov and of total variation and since in this example $S \equiv R$, we can narrow the part on the right side of inequality (4.8) for the next stability inequality:

$$\sup_{x \in [0, U]} \Delta_{V_\alpha}(x) \leq \bar{K}_\alpha \mathbb{V}(D_\xi, D_{\tilde{\xi}}) = \bar{K}_\alpha \int_0^\infty |g_\xi(s) - g_{\tilde{\xi}}(s)| ds, \quad (4.10)$$

where constant \bar{K}_α is given in the inequality (4.9).

❖ **Kantorovich metric (κ).** Let be ($s \geq 0$),

$$G_\xi(s) = \int_0^s g_\xi(z) dz; \quad G_{\tilde{\xi}}(s) = \int_0^s g_{\tilde{\xi}}(z) dz,$$

the distribution functions of random variables ξ and $\tilde{\xi}$, respectively, in Equations (4.1) and (4.2). Then, using the fact that $(\ell \pi)^2 \leq \kappa$, see [18], relates the Lévy-Prokhorov metric and the Kantorovich metric (which was defined in Section 2), the part on the right side of the inequality (4.8) is bounded as

$$\sup_{x \in [0, U]} \Delta_{V_\alpha}(x) \leq \bar{K}_\alpha \left[\kappa(D_\xi, D_{\tilde{\xi}}) \right]^{1/2} = \bar{K}_\alpha \left[\int_0^\infty |G_\xi(s) - G_{\tilde{\xi}}(s)| ds \right]^{1/2}, \quad (4.11)$$

where constant \bar{K}_α is given in the inequality (4.9).

The integral in the last inequality represents the Kantorovich metric between ξ and $\tilde{\xi}$. The inequality (4.11) is more informative compared to inequality (4.10) since it supports that approximation of G_ξ for the corresponding empirical distribution functions.

4.2. Example 4.2

Let be $X = S = \mathbb{R}$, $A(x) = A$, $x \in X$, with A being a compact set in \mathbb{R} . Now, define the following processes:

$$x_t = H(x_{t-1}, a_t) + G(x_{t-1}) \xi_t, \quad t \geq 1;$$

$$\tilde{x}_t = H(\tilde{x}_{t-1}, \tilde{a}_t) + G(\tilde{x}_{t-1}) \tilde{\xi}_t, \quad t \geq 1;$$

where $\xi_t \sim N(\theta, 1)$, $\tilde{\xi}_t \sim N(\tilde{\theta}, 1)$, $H: \mathbb{R} \times A \rightarrow \mathbb{R}$ and $G: \mathbb{R} \rightarrow \mathbb{R}$ are bounded and Lipschitz functions with constants L_H and L_G respectively.

In [19], it is shown that assumption 3.1 is satisfied for this model.

Properly selecting a cost function $c(x, a)$ that is bounded and Lipschitz, it is

assured that the clauses (1) and (2) from assumption 3.2 are fulfilled; for instance, if the following cost function is selected by $c(x, a) = x - a$, then given that $k = (x, a), k' = (x', a') \in \mathbb{K}$, we get that

$$|c(k) - c(k')| = |(x - a) - (x' - a')| \leq 2\nu(k, k'),$$

so, by selecting a constant of $L_o = 2$, this clause (2) is satisfied.

On the other hand, it is clear that the clause (3) is satisfied for any positive constant L_1 . To validate the clause (4) of assumption 3.2 first, define the following random variables:

$$y(k) = F(k, \xi_1) = H(x_{t-1}, a_t) + G(x_{t-1})\xi_1, \quad k = (x_t, a_t) \in \mathbb{K}$$

$$y(k') = F(k', \xi_1) = H(x'_{t-1}, a'_t) + G(x'_{t-1})\xi_1, \quad k' = (x'_t, a'_t) \in \mathbb{K}$$

so, it is clear that the probability density of each of the previous random variables is, respectively

$$y(k) \sim N(H(x_{t-1}, a_t) + \theta G(x_{t-1}), G^2(x_{t-1})) = f_{y(k)},$$

$$y(k') \sim N(H(x'_{t-1}, a'_t) + \theta G(x'_{t-1}), G^2(x'_{t-1})) = f_{y(k')},$$

then, since this example $S = \mathbb{R}$, after some direct calculations we get to the next result

$$\begin{aligned} & \mathbb{V}(F(k, \xi_1), F(k', \xi_1)) \\ &= \mathbb{V}(y(k), y(k')) = \int_{\mathbb{R}} |f_{y(k)}(D) - f_{y(k')}(D)| dD \\ &\leq \sqrt{\frac{2}{\pi}} |H(x_{t-1}, a_t) + \theta G(x_{t-1}) - H(x'_{t-1}, a'_t) - \theta G(x'_{t-1})|, \end{aligned}$$

and as it was assumed that the functions H and G are Lipschitz for the constants L_H, L_G respectively, then from the last inequality we get that

$$\mathbb{V}(F(k, \xi_1), F(k', \xi_1)) \leq \sqrt{\frac{2}{\pi}} (L_H + \theta L_G) \nu(k, k').$$

So, by selecting the constant $L_2 = \sqrt{\frac{2}{\pi}} (L_H + \theta L_G)$ the clause (4) of assumption 3.2 is satisfied. To validate the clause (5) of this assumption, let be $k = (x_t, a_t) \in \mathbb{K}, s, s' \in S$ and note that

$$\rho(F(k, s), F(k, s')) \leq |H(x_{t-1}, a_t) + G(x_{t-1})\xi_t - H(x_{t-1}, a_t) - G(x_{t-1})\xi'_t|,$$

and since the functions H and G are bounded, let M_G be the finite constant, such that $|G(x)| \leq M_G$ for all $x \in X$. Therefore, from the last inequality we get that

$$\rho(F(k, s), F(k, s')) \leq M_G r(s, s').$$

So for a constant of $L_3 = M_G$, the clause (5) is satisfied. Finally, since the function $F(k, s)$ is continuous in all its arguments, then the clause (6) is also true.

In conclusion, the example 4.2 satisfies the assumption 3.2, so the result of the Theorem 3.1 can be applied, see inequalities (3.1), (3.2), and narrow the stability

index using the Lévy-Prokhorov metric

$$\Delta_{V_\alpha} \leq \hat{K} \ell \pi(D_\xi, D_{\bar{\xi}}),$$

where

$$\hat{K} = \frac{4\alpha \left\{ b + M_G(1+L_1) \left[L_0(1-\alpha) + \alpha b \left(\sqrt{\frac{2}{\pi}} (L_H + \theta L_G) \right) \right] \right\}}{(1-\alpha)^3}.$$

5. Proofs

5.1. Some Preliminary Lemmas

For the proof of the theorem 3.1, the following lemmas will be used:

Lemma 5.1. Under assumption 3.2, the value function V_α^* defined in Equation (2.5) satisfies the condition of Lipschitz in the state space X , with the constant $L = (L_1 + 1) \left[L_0 + \frac{\alpha b L_2}{1-\alpha} \right]$.

Proof. For the assumption 3.2 clause (1), for each $\pi \in \Pi$ we get that $V_\alpha(x, \pi) = \mathbb{E}_x \sum_{t=1}^{\infty} \alpha^{t-1} c(x_{t-1}, a_t)$ is bounded by $\frac{b}{1-\alpha}$, then $\|V_\alpha\|_\infty \leq \frac{b}{1-\alpha}$.

On the other hand, in [16] it is proved that the following operators:

$$\begin{aligned} Tu(x) &:= \inf_{a \in A(x)} \left\{ c(x, a) + \alpha \mathbb{E}u[F(x, a, \xi)] \right\} \\ \tilde{T}u(x) &:= \inf_{a \in A(x)} \left\{ c(x, a) + \alpha \mathbb{E}u[F(x, a, \bar{\xi})] \right\} \end{aligned} \quad (5.1)$$

are contractive in the space of Banach \mathbf{B} with module α .

Now, of these operators will be selected the terms that are inside the “brackets” to define the following function:

$$g(x, a) \equiv g(k) := c(k) + \alpha \mathbb{E}V_\alpha[F(k, \xi)], \quad k \in \mathbb{K}.$$

It is claimed that the function $g(k)$ is Lipschitz for the constant $\bar{L} = L_0 + \frac{\alpha b L_2}{1-\alpha}$.

To prove it, let be $k = (x, a)$, $k' = (x', a') \in \mathbb{K}$, then

$$\begin{aligned} |g(k) - g(k')| &= |c(k) + \alpha \mathbb{E}V_\alpha[F(k, \xi)] - c(k') - \alpha \mathbb{E}V_\alpha[F(k', \xi)]| \\ &\leq |c(k) - c(k')| + \alpha |\mathbb{E}V_\alpha[F(k, \xi)] - \mathbb{E}V_\alpha[F(k', \xi)]|. \end{aligned}$$

Applying the assumption 3.2 clause (2) and the fact that $\|V_\alpha\|_\infty \leq \frac{b}{1-\alpha}$, we get that

$$|g(k) - g(k')| \leq L_0 v(k, k') + \alpha \frac{b}{1-\alpha} \mathbb{V}(F(k, \xi), F(k', \xi)).$$

Then, applying the assumption 3.2 clauses (4), the previous inequality can be expressed as

$$\begin{aligned} |g(k) - g(k')| &\leq L_0 v(k, k') + \frac{\alpha b}{1-\alpha} L_2 v(k, k') \\ &= \left[L_0 + \frac{\alpha b L_2}{1-\alpha} \right] v(k, k') = \bar{L} v(k, k'), \end{aligned}$$

therefore, $g(k) \in Lip(\bar{L})$ in \mathbb{K} , where $\bar{L} = \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] v(k, k')$.

By virtue of which the operators given in Equation (5.1) are contractive, we get that $V_\alpha^* = TV_\alpha^* = \inf_{a \in A(x)} \{g(x, a)\}$.

Then, to prove that the function V_α^* is Lipschitz in X with the constant $L = (L_1 + 1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] = (L_1 + 1) \bar{L}$, is enough to try the following:

For the function $g : \mathbb{K} \rightarrow \mathbb{R}$, so that $|g(k) - g(k')| \leq \bar{L}v(k, k')$, with $k, k' \in \mathbb{K}$, it follows that for all $x, y \in X$, $a \in A$:

$$\left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right| \leq L\rho(x, y) = \bar{L}(L_1 + 1)\rho(x, y) \quad (5.2)$$

Remark 5.1. Observe that

$$\begin{aligned} & \left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right| \\ &= \left| \inf_{a \in A(x)} \{c(x, a) + \alpha \mathbb{E}V_\alpha^*[F(x, a, \xi)]\} - \inf_{a \in A(y)} \{c(y, a) + \alpha \mathbb{E}V_\alpha^*[F(y, a, \xi)]\} \right|, \end{aligned}$$

so, we have that

$$\left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right| = |V_\alpha^*(x) - V_\alpha^*(y)|.$$

So if the inequality (5.2) is met, then it is true that

$$|V_\alpha^*(x) - V_\alpha^*(y)| \leq \bar{L}(L_1 + 1)\rho(x, y),$$

which would conclude that V_α^* is Lipschitz in X .

Next the proof of the inequality (5.2) is presented.

Let be $q(x, y) := \left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right|$.

Then, by the inequality of the triangle we get the following:

$$\begin{aligned} q(x, y) &\leq \left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right| + \left| \inf_{a \in A(x)} g(y, a) - \inf_{a \in A(y)} g(y, a) \right|, \\ q(x, y) &\leq \bar{L}\rho(x, y) + I, \end{aligned}$$

where

$$I = \left| \inf_{a \in A(x)} g(y, a) - \inf_{a \in A(y)} g(y, a) \right| \quad (5.3)$$

It will be proven that

$$I \leq \bar{L}L_1\rho(x, y) \quad (5.4)$$

The proof will be done by contradiction. Assuming inequality is not met given in (5.4), then there is a $\varepsilon > 0$ such that the following is satisfied:

$$I > \bar{L}L_1\rho(x, y) + \varepsilon \quad (5.5)$$

Due to the compactness of the sets $A(x)$, $A(y)$ and to the continuity of g , there are elements $a_x \in A(x)$, $a_y \in A(y)$ for which the infimum are reached in I , see Equation (5.3).

If it is admitted for example that

$$I = \inf_{a \in A(x)} g(y, a) - \inf_{a \in A(y)} g(y, a) > \bar{L}L_1\rho(x, y) + \varepsilon,$$

Then

$$\begin{aligned} g(y, a_x) - g(y, a_y) &> \bar{L}L_1\rho(x, y) + \varepsilon, \\ g(y, a_x) &> g(y, a_y) + \bar{L}L_1\rho(x, y) + \varepsilon, \end{aligned} \quad (5.6)$$

Now, for the assumption 3.2 clause (3), as $h(A(x), A(y)) \leq L_1\rho(x, y)$, exists $\bar{a}_x \in A(x)$ such that $\ell(a_y, \bar{a}_x) \leq L_1\rho(x, y)$ and consequently we get that

$$\left| g(y, \bar{a}_x) - g(y, a_y) \right| \leq \bar{L}v(k, k') \leq \bar{L}l(\bar{a}_x, a_y) \leq \bar{L}L_1\rho(x, y).$$

The above implies that

$$\begin{aligned} g(y, \bar{a}_x) &\leq g(y, a_y) + \bar{L}L_1\rho(x, y) \\ g(y, a_y) &\geq g(y, \bar{a}_x) - \bar{L}L_1\rho(x, y), \end{aligned}$$

if this last inequality is substituted in the inequality (5.6), we obtain

$$g(y, a_x) > g(y, \bar{a}_x) + \varepsilon,$$

which contradicts the fact that \bar{a}_x is the element for which the minimum of $g(y, a)$ over $A(x)$ is reached. Therefore, the assumption made in the inequality (5.5) is false. Then, we get that $I \leq \bar{L}L_1\rho(x, y)$, which implies that $q(x, y) \leq \bar{L}\rho(x, y) + \bar{L}L_1\rho(x, y)$ and consequently

$$q(x, y) \in \text{Lip} \left((1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] \right).$$

Finally, because of the comments made in remark 5.1, we get that

$$V_\alpha^* \in \text{Lip} \left((1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] \right),$$

which proves lemma 5.1.

Lemma 5.2. Under assumption 3.2, the value function V_α^* defined in Equation (2.5) satisfies the condition of Lipschitz in space S with the constant

$$L_* = (1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] L_3.$$

Proof. For proof of lemma 5.2, the following function will be used: Let be $k \in \mathbb{K}$; define the function φ_k as $\varphi_k(\cdot) := V_\alpha^*[F(k, \cdot)]: S \rightarrow \mathbb{R}$.

Let be $k \in \mathbb{K}$, $s, s' \in S$. By the definition of functional φ_k we get that

$$\left| \varphi_k(s) - \varphi_k(s') \right| = \left| V_\alpha^*[F(k, s)] - V_\alpha^*[F(k, s')] \right|,$$

and because of lemma 5.1, we came to the next inequality:

$$\left| \varphi_k(s) - \varphi_k(s') \right| \leq (1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] \rho(F(k, s), F(k, s')).$$

Now, applying assumption 3.2 clause (4) to the previous inequality we get the following:

$$|\varphi_k(s) - \varphi_k(s')| \leq (1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] L_3 r(s, s'),$$

namely

$$V_\alpha^* \in \text{Lip} \left((1 + L_1) \left[L_0 + \frac{\alpha b L_2}{1 - \alpha} \right] L_3 \right) \text{ in } S.$$

5.2. The Proof of the Theorem 3.1

To prove inequality (3.1) we take advantage of method proposed in [7], nevertheless we need to modify this technique and the combination of certain Lyapunov like conditions in the results of the paper allows to use the contractive properties of the operators related to the discount cost optimality equations, so the following are the incorporations and developments required for the proof of the Theorem 3.1 obtained in this article, with a bounded cost function.

Let $\pi = \{f, f, \dots\}$, $\tilde{\pi} = \{\tilde{f}, \tilde{f}, \dots\}$ be the optimal stationary policies for processes given in Equations (2.3) and (2.6) respectively, and $V_\alpha^*, \tilde{V}_\alpha^*$ the corresponding value functions.

Then (see chapter 8 of [16]) $V_\alpha^*, \tilde{V}_\alpha^*$ and f, \tilde{f} satisfy the following optimality equations (even more, $V_\alpha^*, \tilde{V}_\alpha^*$ are the only solutions to these equations):

$$\begin{aligned} V_\alpha^*(x) &:= \inf_{a \in \mathcal{A}(x)} \left\{ c(x, a) + \alpha \mathbb{E} V_\alpha^* [F(x, a, \xi)] \right\} \\ &= c(x, f(x)) + \alpha \mathbb{E} V_\alpha^* [F(x, f(x), \xi)], \end{aligned} \quad (5.7)$$

$$\begin{aligned} \tilde{V}_\alpha^*(x) &:= \inf_{a \in \mathcal{A}(x)} \left\{ c(x, a) + \alpha \mathbb{E} \tilde{V}_\alpha^* [F(x, a, \tilde{\xi})] \right\} \\ &= c(x, \tilde{f}(x)) + \alpha \mathbb{E} \tilde{V}_\alpha^* [F(x, \tilde{f}(x), \tilde{\xi})]. \end{aligned} \quad (5.8)$$

For all $(x, a) \in \mathbb{K}$ are defined

$$\begin{aligned} H(x, a) &:= c(x, a) + \alpha \mathbb{E} V_\alpha^* [F(x, a, \xi)], \\ \tilde{H}(x, a) &:= c(x, a) + \alpha \mathbb{E} \tilde{V}_\alpha^* [F(x, a, \tilde{\xi})]. \end{aligned}$$

As it has been proved in [7], the stability index given in Equation (2.8) can be represented as

$$\Delta_{V_\alpha}(x) := \sum_{i=1}^{\infty} \alpha^{i-1} \mathbb{E}_x^{\tilde{\pi}} \Lambda_i \quad (5.9)$$

where

$$\Lambda_t := H(x_{t-1}, a_t) - \inf_{a \in \mathcal{A}(x_{t-1})} H(x_{t-1}, a) \quad (5.10)$$

and $\{x_{t-1}, a_t, t \geq 1\}$ is the trajectory of the process given in Equation (2.3) applying the control policy $\tilde{\pi} = \{\tilde{f}, \tilde{f}, \dots\}$.

By the definition given in Equations (5.10) and (5.8) along with the fact that $\tilde{\pi}$ is optimal for the process given in Equation (2.6), we have that

$$\Lambda_t = H(x_{t-1}, a_t) - \tilde{H}(x_{t-1}, a_t) + \inf_{a \in \mathcal{A}(x_{t-1})} \tilde{H}(x_{t-1}, a) - \inf_{a \in \mathcal{A}(x_{t-1})} H(x_{t-1}, a),$$

which implies

$$|\Lambda_t| \leq 2\alpha \sup_{a \in \mathcal{A}(x_{t-1})} |H(x_{t-1}, a) - \tilde{H}(x_{t-1}, a)|,$$

and by the definition of the functions H and \tilde{H}

$$|\Lambda_t| \leq 2\alpha \sup_{a \in \mathcal{A}(x_{t-1})} \left| \mathbb{E}V_\alpha^* [F(x_{t-1}, a, \xi)] - \mathbb{E}\tilde{V}_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] \right|.$$

Then, applying the inequality of the triangle

$$\begin{aligned} |\Lambda_t| &\leq 2\alpha \sup_{a \in \mathcal{A}(x_{t-1})} \left| \mathbb{E}V_\alpha^* [F(x_{t-1}, a, \xi)] - \mathbb{E}V_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] \right| \\ &\quad + 2\alpha \sup_{a \in \mathcal{A}(x_{t-1})} \left| \mathbb{E}V_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] - \mathbb{E}\tilde{V}_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] \right|. \end{aligned} \quad (5.11)$$

Define now, the next pseudo-metric:

$$\mu(D_\xi, D_{\tilde{\xi}}) := \sup \left\{ \left| \mathbb{E}V_\alpha^* [F(x, a, \xi)] - \mathbb{E}V_\alpha^* [F(x, a, \tilde{\xi})] \right| : (x, a) \in \mathbb{K} \right\} \quad (5.12)$$

Then, from Equation (5.12) it is observed that the first summand on the right side of Equation (5.11) is bounded by $2\alpha\mu(D_\xi, D_{\tilde{\xi}})$.

On the other hand, applying the supremum x_{t-1} in Equation (5.11), the second term on the right side is

$$2\alpha \sup_{a \in \mathcal{A}(x_{t-1})} \left| \mathbb{E}V_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] - \mathbb{E}\tilde{V}_\alpha^* [F(x_{t-1}, a, \tilde{\xi})] \right| \leq 2\alpha \|V_\alpha^* - \tilde{V}_\alpha^*\|_\infty. \quad (5.13)$$

As already mentioned, the operators given in Equation (5.1) are contractive in \mathbf{B} with module α . So, Equations (5.7) and (5.8) can be expressed as $V_\alpha^* = TV_\alpha^*$ and $\tilde{V}_\alpha^* = \tilde{T}\tilde{V}_\alpha^*$. Now, given that V_α^* and \tilde{V}_α^* are fixed points of these operators, we get that

$$\|V_\alpha^* - \tilde{V}_\alpha^*\|_\infty = \|TV_\alpha^* - \tilde{T}\tilde{V}_\alpha^*\|_\infty,$$

now, applying the inequality of the triangle

$$\begin{aligned} \|V_\alpha^* - \tilde{V}_\alpha^*\|_\infty &= \|TV_\alpha^* - \tilde{T}\tilde{V}_\alpha^*\|_\infty + \|\tilde{T}\tilde{V}_\alpha^* - \tilde{T}\tilde{V}_\alpha^*\|_\infty \\ &\leq \|TV_\alpha^* - \tilde{T}\tilde{V}_\alpha^*\|_\infty + \alpha \|V_\alpha^* - \tilde{V}_\alpha^*\|_\infty \\ &\leq \frac{1}{1-\alpha} \|TV_\alpha^* - \tilde{T}\tilde{V}_\alpha^*\|_\infty \\ &\leq \frac{\alpha}{1-\alpha} \sup_{(x,a) \in \mathbb{K}} \left| \mathbb{E}V_\alpha^* [F(x, a, \xi)] - \mathbb{E}\tilde{V}_\alpha^* [F(x, a, \tilde{\xi})] \right|. \end{aligned}$$

Using the definition given in Equation (5.12), we obtain that the previous inequality can be expressed as

$$\|V_\alpha^* - \tilde{V}_\alpha^*\|_\infty \leq \frac{\alpha}{1-\alpha} \mu(D_\xi, D_{\tilde{\xi}}).$$

Substituting this last expression in the inequality (5.13), we get that the second term on the right side of inequality (5.11) is bounded by $\frac{2\alpha}{1-\alpha} \mu(D_\xi, D_{\tilde{\xi}})$ and so inequality (5.11) is bounded by

$$\sup_X |\Lambda_t| \leq 2\alpha\mu(D_\xi, D_{\tilde{\xi}}) + \frac{2\alpha^2}{1-\alpha} \mu(D_\xi, D_{\tilde{\xi}}) = \frac{2\alpha}{1-\alpha} \mu(D_\xi, D_{\tilde{\xi}}). \quad (5.14)$$

Finally, substituting inequality (5.14) in Equation (5.9) we obtain

$$\sup_x \Delta_{V_\alpha}(x) \leq \sum_{i=1}^{\infty} \alpha^{i-1} \left\{ \frac{2\alpha}{1-\alpha} \mu(D_\xi, D_{\bar{\xi}}) \right\} = \frac{2\alpha}{(1-\alpha)^2} \mu(D_\xi, D_{\bar{\xi}}). \quad (5.15)$$

To find a dimension for $\mu(D_\xi, D_{\bar{\xi}})$, it will be used the definition of the Dudley metric (δ) in the space of distributions in (S, r) :

$$\delta(D_\xi, D_{\bar{\xi}}) = \delta(P_X, P_Y) := \sup_{\varphi \in D} |\mathbb{E}\varphi(X) - \mathbb{E}\varphi(Y)|,$$

where

$$D := \{ \varphi : S \rightarrow \mathbb{R} : \|\varphi\|_\infty + \|\varphi\|_L \leq 1 \},$$

and

$$\|\varphi\|_\infty = \sup_{x \in S} |\varphi(x)|; \quad \|\varphi\|_L = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{r(x, y)}.$$

(See [15] for definition and properties of δ).

By the lemma 5.2, we get that $V_\alpha^* \in Lip\left((1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]L_3\right)$ and since $\|V_\alpha^*\|_\infty \leq \frac{b}{1-\alpha}$, then the stability index can be narrowed in terms of Dudley's metric by the following expression:

$$\Delta_{V_\alpha}(x) = \frac{2\alpha}{(1-\alpha)^2} \left[\frac{b}{1-\alpha} + (1+L_1) \left[L_0 + \frac{\alpha b L_2}{1-\alpha} \right] L_3 \right] \delta(D_\xi, D_{\bar{\xi}}).$$

Now, using the well-known relationship $\delta \leq 2\ell\pi$ between the Dudley metric and Lévy-Prokhorov metric (see [18]) and after some direct calculations, the desired inequality (3.1) is obtained with the constant given in Equation (3.2).

6. Conclusions

Despite the vast literature that exists on the subject of Markov controllable processes, a few studies have been carried out on the subject of estimating stability. The study of stability for Markov control processes represents a challenge both from a theoretical and a practical point of view. Proposing appropriate probabilistic metrics to achieve so-called stability inequalities is an additional effort.

In this article, conditions were found to obtain the stability of a Markov control process under the optimization criterion of expected total α -discounted cost with a bounded cost function using the Lévy-Prokhorov metric.

The importance of being able to use the Lévy-Prokhorov metric lies in the fact that for application problems it allows estimations of the stability index under the use of empirical distributions for the random elements, since they converge weakly under this metric to the distributions that are it tries to estimate (unlike the so-called "strong metrics").

On the other hand, since in applications, there is no company that can bear unlimited (unbounded) costs, the results found in this work using simple tech-

niques such as contractive operators provide an estimate of the increase in cost (the stability index) to control the “original process” using the optimal policy of the “approximate process”. Of course, the stability constant (\bar{K}_α) affects this stability index, specifically in this work it was found that this constant is of order $O(1-\alpha)^{-3}$ if $\alpha \rightarrow 1$. There are arguments to support the hypothesis that in the left part of the inequality (3.1) (for each initial state fixed x): $\Delta_{V_\alpha} \rightarrow \infty$ when $\alpha \rightarrow 1$ and the distribution of ξ and $\tilde{\xi}$ are fixed. It is not clear what the rate of such growth is. Therefore, it is proposed that in future research, based on particular (and simple) control processes, to verify the growth rate of Δ_{V_α} using computational experiments and process simulation.

Acknowledgements

The author is particularly grateful to Professor Edgar Vladyvosky M.S. for his instructive discussions on a generalization of the Markov processes and properties of the Lévy-Prokhorov metric.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Dynkin, E.B. and Yushkevich, A.A. (1979) Controlled Markov Processes. Springer-Verlag, New York. <https://doi.org/10.1007/978-1-4615-6746-2>
- [2] Hernandez-Lerma, O. (1989) Adaptive Markov Control Process. Springer-Verlang, New York. <https://doi.org/10.1007/978-1-4419-8714-3>
- [3] Gordienko, E.I. (1988) Stability Estimates for Controlled Markov Chains with a Minorant. *Journal of Soviet Mathematics*, **40**, 481-486. <https://doi.org/10.1007/BF01083641>
- [4] Gordienko, E.I. (1992) An Estimate of the Stability of Optimal Control of Certain in Stochastic and Deterministic Systems. *Journal of Soviet Mathematics*, **59**, 891-899. <https://doi.org/10.1007/BF01099115>
- [5] Gordienko, E.I. and Salem, F.S. (1998) Robustness Inequality for Markov Control Processes with Undounded Costs. *Systems & Control Letters*, **33**, 125-130. [https://doi.org/10.1016/S0167-6911\(97\)00077-7](https://doi.org/10.1016/S0167-6911(97)00077-7)
- [6] Gordienko, E.I. and Yushkevich, A. (2003) Stability Estimates in the Problem of Optimal Switching of a Markov Chain. *Mathematical Methods of Operations Research*, **57**, 345-365. <https://doi.org/10.1007/s001860200258>
- [7] Gordienko, E.I., Lemus-Rodriguez, E. and Montes-de-Oca, R. (2008) Discounted Cost Optimality Problem: Stability with Respect to Weak Metrics. *Mathematical Methods of Operations Research*, **68**, 77-96. <https://doi.org/10.1007/s00186-007-0171-z>
- [8] Rachev, S.T. and Ruschendorf, L. (1998) Mass Transportation Problem, Vol. II: Applications. Springer, New York.
- [9] Gordienko, E.I., Lemus-Rodriguez, E. and Montes-de-Oca, R. (2009) Average Cost Markov Control Processes: Stability with Respect to the Kantorovich Metric. *Mathematical Methods of Operations Research*, **70**, 13-33.

- <https://doi.org/10.1007/s00186-008-0229-6>
- [10] Montes-de-Oca, R. and Salem-Silva, F. (2005) Estimates for Perturbations of Average Markov Decision Processes with a Minimal State and Upper Bounded by Stochastically Ordered Markov Chains. *Kybernetika*, **41**, 757-772.
- [11] Aziz, M.M. and Merie, D.M. (2020) Stability and Adaptive Control with Synchronization of 3-D Dynamical System. *Open Access Library Journal*, **7**, e6075.
<https://doi.org/10.4236/oalib.1106075>
- [12] Yilmaz, S., Büyükköroglu, T. and Dzhafarov, V. (2015) On Asymptotic Stability of Linear Control Systems. *Applied Mathematics*, **6**, 71-77.
<https://doi.org/10.4236/am.2015.61008>
- [13] Engblom, S. (2014) On the Stability of Stochastic Jump Kinetics. *Applied Mathematics*, **5**, 3217-3239. <https://doi.org/10.4236/am.2014.519300>
- [14] Hernandez-Lerma, O. and Lasserre, J.B. (1996) Discrete-Time Markov Control Processes. Basic Optimality Criteria. Springer, New York.
<https://doi.org/10.1007/978-1-4612-0729-0>
- [15] Rachev S. T. (1991) Probability Metrics and the Stability of Stochastic Models. Wiley, Chichester.
- [16] Hernandez-Lerma O. and Lasserre J. B. (1999) Further Topics on Discrete-Time Markov Control Processes. Springer, New York.
<https://doi.org/10.1007/978-1-4612-0561-6>
- [17] Yakowitz, S. (1982) Dynamic Programming Applications in Water Resources. *Water Resources Research*, **18**, 673-696. <https://doi.org/10.1029/WR018i004p00673>
- [18] Dudley, R.M. (1989) Real Analysis and Probability. Wadsworth & Brooks Cole, Pacific Grove.
- [19] Arapostathis, A., Borkar, V.S., Fernandez-Gaucherand, G.M.K. and Marcus, S.I. (1993) Discrete-Time Controlled Markov Processes with Average Cost Criterion: A Survey. *SIAM Journal on Control and Optimization*, **31**, 282-344.
<https://doi.org/10.1137/0331018>