LP Based Bounds for Cesàro and Abel Limits of the Optimal Values in Non-ergodic Stochastic Systems

Konstantin Avrachenkov[1], Vladimir Gaitsgory[2] and Lucas Gamertsfelder[2]

***Abstract*— In this paper, we study asymptotic properties of problems of control of stochastic discrete time systems with time averaging and time discounting optimality criteria, and we establish that the Cesàro and Abel limits of the optimal values in such problems can be estimated with the help of a certain infinite-dimensional (ID) linear programming (LP) problem and its dual.**

I. INTRODUCTION

In this paper, we study asymptotic properties of problems of control of stochastic discrete time systems with time averaging and time discounting optimality criteria, and we establish that the Cesàro and Abel limits of the optimal values in such problems can be evaluated with the help of a certain infinite-dimensional (ID) linear programming (LP) problem and its dual. Note that matters related to the existence and the equality of such limits have been investigated by many; see, e.g., [1], [2], [6], [8], [14], [18], [21], [22], [30], [33], [34], [37], [41]. A distinct feature of the present paper is that the Cesàro and Abel limits of the optimal values are evaluated with the help of LP tools.

Allowing one to use convex duality theory, the LP approach to various classes of optimal control problems have been studied extensively in the literature. For example, LP formulations of problems of optimal control of stochastic systems evolving in continuous time have been considered in [2], [9], [13], [15], [31], [39]. Various aspects of the LP approach to problems of optimization of discrete time stochastic systems (controlled Markov chains) or continuous time systems with random jumps have been discussed in [10], [23]-[29], [35], [36]. In the deterministic setting, the LP approach has been developed/applied in [20], [24], [32], [38], [42] for systems evolving in continuous time considered on a finite time interval. The applicability of the LP approach to deterministic continuous and discrete time systems considered on the infinite time horizon has been explored in [11], [12], [17]-[19].

In [11] and [12], in particular, it was shown (for deterministic, continuous time systems in [11], and for deterministic, discrete time systems in [12]) that the Cesàro and Abel limits of the optimal values are bounded from above by optimal values of certain infinite-dimensional linear programming (IDLP) problems and that these limits are bounded from below by the optimal values of the corresponding dual problems. This paper extends this line of research to stochastic systems evolving in discrete time. In more detail, we consider the discrete time stochastic control system

$$y(t+1) = f(y(t), u(t), s(t)), \quad t = 0, 1, 2, \dots, \tag{1}$$

and we assume that the following conditions are satisfied:

- The function $f(y,u,s) : Y \times \hat{U} \times S \to \mathbb{R}^m$ is continuous in (y,u) on $Y \times \hat{U}$ and is Borel measurable in s on S, where Y is a compact subset of $\mathbb{R}^n$, $\hat{U}$ is a compact metric space, and S is a Polish space.
- $s(t) \in S,\ t = 0, 1, ...,$ is a sequence of independent, identically distributed random elements.
- The controls $u(t),\ t = 0, 1, ...,$ are defined by a sequence of functions $\pi \stackrel{\text{def}}{=} \{\pi_t(y),\ t = 0, 1, ...\}$ that are Borel measurable selections of a multivalued map $U(\cdot) : Y \rightsquigarrow \hat{U}$ so that
$$u(t) = \pi_t(y(t)) \in U(y(t)), \quad t = 0, 1, ..., \tag{2}$$
where $U(\cdot)$ is upper semicontinuous and compact-valued (that is, $U(y)$ is compact for any $y \in Y$).
- $f(y,u,s) \in Y$ for any $y \in Y$, any $u \in U(y)$, and any $s \in S$ (that is, the set Y is forward invariant with respect to system (1)).

Let Π stand for the set of sequences of measurable selections of $U(\cdot)$:

$$\Pi \stackrel{\text{def}}{=} \{\pi = \{\pi_t(\cdot),\ t = 0, 1, ...\} \mid \pi_t(y) \in U(y)\ \forall\ y \in Y,$$
$$\pi_t(\cdot) \text{ are Borel measurable}\}.$$

For any sequence $\pi \in \Pi$ (for convenience, such sequences will be referred to as *control plans*) and any initial condition $y(0) = y_0 \in Y$, let $(y^{\pi,y_0}(\cdot), u^{\pi,y_0}(\cdot))$ stand for the state-control trajectory obtained in accordance with (1) and (2).

Consider the following optimal control problems

$$\frac{1}{T} \min_{\pi \in \Pi} E\left[\sum_{t=0}^{T-1} k(y^{\pi,y_0}(t), u^{\pi,y_0}(t))\right] \stackrel{\text{def}}{=} v_T(y_0), \tag{3}$$

$$\epsilon \min_{\pi \in \Pi} \sum_{t=0}^{\infty} (1-\epsilon)^t k(y^{\pi,y_0}(t), u^{\pi,y_0}(t)) \stackrel{\text{def}}{=} h_\epsilon(y_0), \tag{4}$$

where $k(y,u) : Y \times \hat{U} \to \mathbb{R}$ is a continuous function with

$$|k(y,u)| \le M \quad \forall (y,u) \in Y \times \hat{U}, \qquad M = \text{const},$$

and $\epsilon \in (0,1)$ (that is, $(1-\epsilon)$ is a discount factor).

[1]K. Avrachenkov is with INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 Valbonne, France. k.avrachenkov@inria.fr

[2]V. Gaitsgory and L. Gamertsfelder are with Department of Mathematics and Statistics, Macquarie University, Eastern Road, Macquarie Park, NSW 2113, Australia. vladimir.gaitsgory@mq.edu.au, lucas.gamertsfelder@hdr.mq.edu.au L. Gamertsfelder is the corresponding author. This is the author version of the paper accepted at European Control Conference 2021 (ECC21).

Proposition 1.1: Under the assumptions made above, the optimal value functions $v_T(\cdot)$ and $h_\epsilon(\cdot)$ are lower semicontinuous and satisfy the equations:

$$Tv_T(y) = \min_{u\in U(y)}\{k(y,u)+(T-1)E[v_{T-1}(f(y,u,s))]\}, \quad (5)$$

$$h_\epsilon(y) = \min_{u\in U(y)}\{\epsilon k(y,u)+(1-\epsilon)E[h_\epsilon(f(y,u,s))]\} \ \forall\, y \in Y. \quad (6)$$

Also, the minima in (3) and (4) are achieved.

Proof. The proof follows Theorems 2.4.6 and 7.2.1 in [7]. □

We will be interested in evaluating $\lim_{T\to\infty} v_T(y_0)$ and $\lim_{\epsilon\to 0} h_\epsilon(y_0)$ (these limits are commonly referred to as *Cesàro* and *Abel* limits). More specifically, we will establish that $\limsup_{T\to\infty} v_T(y_0)$ and $\limsup_{\epsilon\to 0} h_\epsilon(y_0)$ are bounded from above by the optimal value of a certain IDLP problem, and that $\liminf_{T\to\infty} v_T(y_0)$ and $\liminf_{\epsilon\to 0} h_\epsilon(y_0)$ are bounded from below by the optimal value of the corresponding dual problem (see Theorem 3.5). An immediate consequence of this result is the statement that the Cesàro and Abel limits of the optimal values exist and are equal to each other if there is no duality gap (see Corollary 3.6).

Note that many results obtained in this paper are stated and proved similarly to their deterministic counterparts obtained in [12]. However, the part of Theorem 3.5 that establishes the upper bound for the Cesàro and Abel limits of the optimal values (this being the most important result of the paper) is stronger than the corresponding statement (Theorem 3.1) in [12] since, in contrast to the latter, it is not assumed that $v_T(\cdot)$ and $h_\epsilon(\cdot)$ are continuous.

For the convenience of the reader, we leave out most of the proofs from the following sections. These proofs may be found in the preprint this work is based on (see [5]).

II. MODEL

Let us introduce some notations and definitions that will be used in the subsequent sections. Let G stand for the graph of $U(\cdot)$,

$$G \stackrel{\text{def}}{=} \text{graph}(U) = \{(y,u) : u \in U(y),\ y \in Y\},$$

and let $\mathcal{P}(G)$ stand for the set of probability measures defined on Borel subsets of G. (Note that, due to upper semicontinuity of $U(\cdot)$, the graph G is a compact subset of $Y \times \hat{U}$.) Given a control plan $\pi \in \Pi$ and an initial condition $y(0) = y_0 \in Y$, denote by $\gamma^{\pi,y_0,T} \in \mathcal{P}(G)$ and $\gamma_d^{\pi,y_0,\epsilon} \in \mathcal{P}(G)$ the probability measures defined as follows: for any Borel $Q \subset G$,

$$\gamma^{\pi,y_0,T}(Q) = \frac{1}{T}E\left[\sum_{t=0}^{T-1} 1_Q(y^{\pi,y_0}(t), u^{\pi,y_0}(t))\right], \quad (7)$$

$$\gamma_d^{\pi,y_0,\epsilon}(Q) = \epsilon E\left[\sum_{t=0}^{\infty}(1-\epsilon)^t 1_Q(y^{\pi,y_0}(t), u^{\pi,y_0}(t))\right], \quad (8)$$

where $1_Q(\cdot)$ is the indicator function of Q. The measures defined by (7) and (8) will be referred to as *occupational measure* and, respectively, *discounted occupational measure* generated by the control plan π. Note that from (7) and (8) it follows that

$$\int_G q(y,u)\gamma^{\pi,y_0,T}(dy,du) = \frac{1}{T}E\left[\sum_{t=0}^{T-1} q(y^{\pi,y_0}(t), u^{\pi,y_0}(t))\right] \quad (9)$$

and

$$\int_G q(y,u)\gamma_d^{\pi,y_0,\epsilon}(dy,du) = \epsilon E\left[\sum_{t=0}^{\infty}(1-\epsilon)^t q(y^{\pi,y_0}(t), u^{\pi,y_0}(t))\right] \quad (10)$$

for any Borel measurable function q on G. In fact, the definitions (7) and (8) are equivalent to that the equality (9) and (10) are valid if $q(\cdot)$ is an indicator function of Q. Therefore, these equalities are valid for linear combinations of indicator functions. The validity of (9) and (10) for any Borel function follows from the fact that any such function can be presented as uniform limit of linear combinations of indicator functions.

Let us denote by $\Gamma_T(y_0)$ the set of occupational measures and by $\Theta_\epsilon(y_0)$ the set of discounted occupational measures:

$$\Gamma_T(y_0) \stackrel{\text{def}}{=} \bigcup_{\pi\in\Pi}\{\gamma^{\pi,y_0,T}\}, \qquad \Theta_\epsilon(y_0) \stackrel{\text{def}}{=} \bigcup_{\pi\in\Pi}\{\gamma_d^{\pi,y_0,\epsilon}\}.$$

Note that, due to (9) and (10), problems (3) and (4) can be rewritten in the form

$$\min_{\gamma\in\Gamma_T(y_0)} \int_G k(y,u)\gamma(dy,du) = v_T(y_0) \quad (11)$$

and

$$\min_{\gamma\in\Theta_\epsilon(y_0)} \int_G k(y,u)\gamma(dy,du) = h_\epsilon(y_0), \quad (12)$$

respectively.

To describe convergence properties of occupational measures, the following metric on $\mathcal{P}(G)$ will be used:

$$\rho(\gamma',\gamma'') := \sum_{j=1}^{\infty}\frac{1}{2^j}\left|\int_G q_j(y,u)\gamma'(dy,du) - \int_G q_j(y,u)\gamma''(dy,du)\right|$$

for $\gamma', \gamma'' \in \mathcal{P}(G)$, where $q_j(\cdot)$, $j = 1,2,\ldots$, is a sequence of Lipschitz continuous functions dense in the unit ball of the space of continuous functions $C(G)$ from G to $\mathbb{R}$. This metric is consistent with the weak* convergence topology on $\mathcal{P}(G)$, that is, a sequence $\gamma^k \in \mathcal{P}(G)$ converges to $\gamma \in \mathcal{P}(G)$ in this metric if and only if

$$\lim_{k\to\infty}\int_G q(y,u)\gamma^k(dy,du) = \int_G q(y,u)\gamma(dy,du)$$

for any $q \in C(G)$. Using the metric ρ, we can define the "distance" $\rho(\gamma,\Gamma)$ between $\gamma \in \mathcal{P}(G)$ and $\Gamma \subset \mathcal{P}(G)$ and the Hausdorff metric $\rho_H(\Gamma_1, \Gamma_2)$ between $\Gamma_1 \subset \mathcal{P}(G)$ and $\Gamma_2 \subset \mathcal{P}(G)$ as follows:

$$\rho(\gamma, \Gamma) \stackrel{\text{def}}{=} \inf_{\gamma' \in \Gamma} \rho(\gamma, \gamma'),$$

$$\rho_H(\Gamma_1, \Gamma_2) \stackrel{\text{def}}{=} \max\{\sup_{\gamma \in \Gamma_1} \rho(\gamma, \Gamma_2), \sup_{\gamma \in \Gamma_2} \rho(\gamma, \Gamma_1)\}.$$

Note that, although, by some abuse of terminology, we refer to $\rho_H(\cdot, \cdot)$ as a metric on the set of subsets of $\mathcal{P}(G)$, it is, in fact, a semi-metric on this set (since $\rho_H(\Gamma_1, \Gamma_2) = 0$ implies $\Gamma_1 = \Gamma_2$ if Γ_1 and Γ_2 are closed, but the equality may not be true if at least one of these sets is not closed).

III. MAIN RESULTS

Consider the IDLP problem

$$\min_{\gamma \in W} \int_G k(y, u)\gamma(dy, du) \stackrel{\text{def}}{=} k^* \tag{13}$$

where W be defined by the equation

$$W \stackrel{\text{def}}{=} \{\gamma \in \mathcal{P}(G) |$$

$$\int_G \big(E\big[\varphi(f(y, u, s))\big] - \varphi(y)\big)\gamma(dy, du) = 0 \quad \forall \varphi \in C(Y)\}.$$

Proposition 3.1: The following equalities are valid:

$$\lim_{\epsilon \to 0} \rho_H(\bar{\text{co}}\Theta_\epsilon, W) = 0, \quad \text{where} \quad \Theta_\epsilon \stackrel{\text{def}}{=} \bigcup_{y_0 \in Y} \{\Theta_\epsilon(y_0)\},$$

$$\lim_{T \to \infty} \rho_H(\bar{\text{co}}\Gamma_T, W) = 0, \quad \text{where} \quad \Gamma_T \stackrel{\text{def}}{=} \bigcup_{y_0 \in Y} \{\Gamma_T(y_0)\}.$$

where $\bar{\text{co}}$ stands for the closed convex hulls of the corresponding sets.

Proof. The proof is given in [5]. □

Corollary 3.2: Due to (11) and (12), from Proposition 3.1 it follows that:

$$\lim_{\epsilon \to 0} \min_{y_0 \in Y} h_\epsilon(y_0) = k^*, \tag{14}$$

$$\lim_{T \to \infty} \min_{y_0 \in Y} v_T(y_0) = k^*. \tag{15}$$

Proposition 3.3: The set W allows the following (equivalent) representation:

$$W = \left\{\gamma \in \mathcal{P}(G) \middle| \gamma_1(Q) = \int_G P(Q|y, u)\gamma(dy, du) \right. \tag{16}$$

$$\left. \forall \text{ Borel } Q \subset Y \right\},$$

where γ_1 is the marginal of γ, that is,

$$\gamma_1(Q) = \int_G 1_Q(y)\gamma(dy, du),$$

and $P(dy|y, u)$ is the transition law associated with system (1), that is,

$$P(Q|y, u) = E[1_Q(f(y, u, s))] \quad \forall\, (y, u) \in G.$$

Proof. The proof is given in [5]. □

REMARK. The validity of the representation (16) makes the results established by Proposition 3.1 and Corollary 3.2 consistent with well known results in Markov control processes theory; see [23]-[27], [35], [36] and references therein. Many of the latter are obtained under assumptions that are lighter than the assumptions we are using in this paper. Note that some of our assumptions can be relaxed too. For example, the assumption about compactness of the state space Y can be replaced by the assumption about the tightness of the set of occupational measures that make the results of Proposition 3.1 valid. However, to make the presentation more expository, we stick to using simpler (albeit more restrictive) assumptions.

Let us consider the following IDLP problem

$$\inf_{(\gamma, \xi) \in \Omega(y_0)} \int_G k(y, u)\gamma(dy, du) \stackrel{\text{def}}{=} k^*(y_0), \tag{17}$$

where the feasible set $\Omega(y_0)$ is defined by $(\gamma, \xi) \in \mathcal{P}(G) \times \mathcal{M}_+(G)$ such that

$$\Omega(y_0) \stackrel{\text{def}}{=} \{(\gamma, \xi) \in \mathcal{P}(G) \times \mathcal{M}_+(G) | \, \gamma \in W,$$

$$\int_G (\varphi(y_0) - \varphi(y))\gamma(dy, du) +$$

$$\int_G \big(E\big[\varphi(f(y, u, s))\big] - \varphi(y)\big)\xi(dy, du) = 0$$

$$\forall \varphi \in C(Y)\},$$

where $\mathcal{M}_+(G)$ stands for the space of nonnegative measures defined on Borel subsets of G. One way to obtain the above problem is to follow formally the construction of [3] and [4]; it is obtained, in fact, by augmenting problem (13) with additional constraints and an additional "decision variable" ξ.

The problem dual to the augmented IDLP problem (17) can be written in the following form

$$\sup_{(\mu, \psi, \eta) \in \mathcal{D}(y_0)} \mu \stackrel{\text{def}}{=} d^*(y_0), \tag{18}$$

where $\mathcal{D}(y_0)$ is the set of triplets $(\mu, \psi(\cdot), \eta(\cdot)) \in I\!R \times C(Y) \times C(Y)$ that for all $(y, u) \in G$ satisfy the inequalities

$$\begin{aligned} & k(y, u) + \\ & (\psi(y_0) - \psi(y)) + E[\eta(f(y, u, s))] - \eta(y) - \mu \geq 0, \\ & E[\psi(f(y, u, s))] - \psi(y) \geq 0. \end{aligned} \tag{19}$$

Note that the optimal value of problem (18) can be equivalently represented as

$$d^*(y_0) = \sup_{(\psi, \eta) \in C(Y) \times C(Y)} \min_{(y, u) \in G} \{k(y, u) +$$

$$(\psi(y_0) - \psi(y)) + E[\eta(f(y, u, s))] - \eta(y)\}, \tag{20}$$

where ψ satisfies the second inequality in (19).

The following proposition establishes the validity of the *weak duality* inequality.

Proposition 3.4: The optimal values of (17) and (18) are related by the inequality

$$d^*(y_0) \le k^*(y_0). \tag{21}$$

Proof. Take any $(\gamma, \xi) \in \Omega(y_0)$ and $(\mu, \psi, \eta) \in \mathcal{D}(y_0)$. Integrating the first inequality in (19) with respect to γ and taking into account that $\gamma \in W$, we conclude that

$$\int_G k(y,u)\gamma(dy,du) + \int_G (\psi(y_0) - \psi(y))\gamma(dy,du) \ge \mu.$$

Since $(\gamma, \xi) \in \Omega(y_0)$, from the second inequality in (19) it follows that

$$\int_G (\psi(y_0) - \psi(y))\gamma(dy,du) =$$

$$-\int_G \big(E[\psi(f(y,u,s))] - \psi(y)\big)\xi(dy,du) \le 0.$$

Therefore,

$$\int_G k(y,u)\gamma(dy,du) \ge \mu.$$

Taking first inf over all $(\gamma,\xi) \in \Omega(y_0)$ in the left-hand-side and then sup over all $(\mu,\psi,\eta) \in \mathcal{D}(y_0)$ in the right-hand-side, one establishes the validity of (21). □

Our main result is the following theorem.

Theorem 3.5: The lower and upper Cesàro/Abel limits of the optimal value functions in problems (3) and (4) satisfy the inequalities, $\forall\ y_0 \in Y$,

$$d^*(y_0) \le \liminf_{T\to\infty} v_T(y_0) \le \limsup_{T\to\infty} v_T(y_0) \le k^*(y_0),$$

$$d^*(y_0) \le \liminf_{\epsilon\to 0} h_\epsilon(y_0) \le \limsup_{\epsilon\to 0} h_\epsilon(y_0) \le k^*(y_0),$$

where $k^*(y_0)$ is the optimal value of the augmented IDLP problem (17) and $d^*(y_0)$ is the optimal value of its dual (18).
Proof. The proof is given in [5]. □

We have also the following immediate corollary.

Corollary 3.6: Let, for a given $y_0 \in Y$, the strong duality equality be valid:

$$k^*(y_0) = d^*(y_0). \tag{22}$$

Then the Cesàro and Abel limits of the optimal values exist and are equal:

$$\lim_{T\to\infty} v_T(y_0) = \lim_{\epsilon\to 0} h_\epsilon(y_0) = k^*(y_0) = d^*(y_0). \tag{23}$$

IV. EXAMPLE

Let the dynamics be one-dimensional and be described by the equation (compare with (1))

$$y(t+1) = y(t)u(t)s(t) \quad \forall\ t = 0, 1, \dots\ , \tag{24}$$

where $Y = [-1,1]$ and $U(y) = \{-1,1\}$ (that is, the control can be either equal to 1 or to -1). Assume that $s(t)$ takes only two values: $s(t) = 1$ with probability $\frac{3}{4}$ and $s(t) = -1$ with probability $\frac{1}{4}$. Consider problems (3) and (4) with $k(y,u) = y$. It can be readily understood, that, in this example, the plan $\pi^* = \{\pi_t^*(y),\ t = 0,1,...\}$, where, for any $t = 0,1,....,$

$$\pi_t^*(y) = +1 \ \text{ for } \ y \in [-1,0] \text{ and } \pi_t^*(y) = -1 \text{ for } y \in (0,1],$$

is optimal in both problem (3) and problem (4) (as this is the plan that maximizes the probability for the state variable to be negative). The optimal values of problems (3) and (4) can be evaluated to be as follows

$$v_T(y_0) = -\frac{1}{2}|y_0| + \frac{1}{T}\left(y_0 + \frac{1}{2}|y_0|\right) \quad \forall\ y_0 \in Y, \tag{25}$$

$$h_\epsilon(y_0) = -\frac{1}{2}|y_0| + \epsilon\left(y_0 + \frac{1}{2}|y_0|\right) \quad \forall\ y_0 \in Y. \tag{26}$$

(By a direct substitution, one can verify that $v_T(y_0)$ and $h_\epsilon(y_0)$, defined in accordance with (25) and (26), satisfy the dynamic programming equations (5) and (6), respectively.) From (25) and (26) it follows that

$$\lim_{\epsilon\to 0} h_\epsilon(y_0) = \lim_{T\to\infty} v_T(y_0) = -\frac{1}{2}|y_0|. \tag{27}$$

The augmented IDLP problem (17) takes the form

$$\inf_{(\gamma,\xi)\in\Omega(y_0)} \int_G y\gamma(dy,du) = k^*(y_0), \tag{28}$$

where $\Omega(y_0)$ is the set of pairs $(\gamma,\xi) \in \mathcal{P}(G) \times \mathcal{M}_+(G)$ that satisfy the equations

$$\int_G \left(\frac{3}{4}\varphi(yu) + \frac{1}{4}\varphi(-yu) - \varphi(y)\right)\gamma(dy,du) = 0, \tag{29}$$

$\forall\ \varphi \in C([-1,1])$, and

$$\int_G (\varphi(y_0) - \varphi(y))\gamma(dy,du) = \tag{30}$$

$$-\int_G \left(\frac{3}{4}\varphi(yu) + \frac{1}{4}\varphi(-yu) - \varphi(y)\right)\xi(dy,du),$$

$\forall\,\varphi \in C([-1,1])$, and where $G = Y \times U = [-1,1] \times \{-1,1\}$ in this case. The corresponding dual problem (see (20)) is

$$\sup_{(\psi,\eta)\in C([-1,1])\times C([-1,1])} \min_{(y,u)\in G} \{y + (\psi(y_0) - \psi(y)) \tag{31}$$

$$+\frac{3}{4}\eta(yu)) + \frac{1}{4}\eta(-yu) - \eta(y)\} = d^*(y_0),$$

where the ψ functions are assumed to satisfy the inequality

$$\frac{3}{4}\psi(yu) + \frac{1}{4}\psi(-yu) - \psi(y) \ge 0, \tag{32}$$

$\forall\ y \in [-1,1]$, and $\forall\ u \in \{-1,1\}$. If a function φ is even, then $\frac{3}{4}\varphi(yu) + \frac{1}{4}\varphi(-yu) - \varphi(y) \equiv 0$ (since u is either equal to 1 or to -1). Therefore, (29) is satisfied for all $\gamma \in \mathcal{P}(G)$, while (30) is converted to $\int_G (\varphi(y_0) - \varphi(y))\gamma(dy,du) = 0$ in this case. The latter equality implies that $\int_G |y_0|^l \gamma(dy,du) = \int_G |y|^l \gamma(dy,du)$ for any $l = 1,2,...$, which, in turn, implies that

$$\gamma(Y_{y_0}) = 1, \quad \text{where} \quad Y_{y_0} \stackrel{\text{def}}{=} \{y : |y| = |y_0|\}.$$

Thus, the constraints (30) ensure that the occupational measures γ generated by the state-control trajectories satisfy the property $\gamma(Y \setminus Y_{y_0}) = 0$. This is consistent with the system's dynamics (see (24)), according to which the only states attended by the state trajectories are y_0 and $-y_0$.

Let

$$\bar{\gamma}(dy,du) \stackrel{\text{def}}{=} \left(\frac{3}{4}\delta_{-|y_0|}(dy) + \frac{1}{4}\delta_{|y_0|}(dy) \right) \delta_{\kappa(y)}(du),$$

$$\bar{\xi}(dy,du) \stackrel{\text{def}}{=} \delta_{y_0}(dy)\delta_{\kappa(y)}(du),$$

where δ_a stands for the Dirac measure concentrated at a, and where $\kappa(y)$ is equal to 1 for $y \in [-1,0]$ and equal to -1 for $y \in (0,1]$ (that is, for an arbitrary function $q(u)$ on U, $\int_U q(u)\delta_{\kappa(y)}(du) = q(1)\ \forall\ y \in [-1,0]$ and $\int_U q(u)\delta_{\kappa(y)}(du) = q(-1)\ \forall\ y \in (0,1]$).

Via a direct substitution into (29) and (30), it can be verified that $(\bar{\gamma},\bar{\xi}) \in \Omega(y_0)$ (note that it is sufficient to verify the validity of (29) and (30) only for the even and odd test functions $\varphi(\cdot)$). Therefore,

$$k^*(y_0) \leq \int_G y\bar{\gamma}(dy,du) = -\frac{1}{2}|y_0|. \tag{33}$$

On the other hand, it can also be verified that the pair of functions $(\bar{\psi}(y), \bar{\eta}(y))$,

$$\bar{\psi}(y) \stackrel{\text{def}}{=} -\frac{1}{2}|y|, \qquad \bar{\eta}(y) \stackrel{\text{def}}{=} \left(y + \frac{1}{2}|y| \right),$$

satisfy the relationships

$$\min_{y\in[-1,1]} \min_{u\in\{-1,1\}} \{y + (\bar{\psi}(y_0) - \bar{\psi}(y))$$
$$+\frac{3}{4}\bar{\eta}(yu)) + \frac{1}{4}\bar{\eta}(-yu) - \bar{\eta}(y)\} = -\frac{1}{2}|y_0|,$$
$$\frac{3}{4}\bar{\psi}(yu)) + \frac{1}{4}\bar{\psi}(-yu) - \bar{\psi}(y) = 0,$$

$\forall\ y \in [-1,1]$, and $\forall\ u \in \{-1,1\}$. Therefore (compare the latter with (31) and (32)),

$$-\frac{1}{2}|y_0| \leq d^*(y_0).$$

This inequality, along with (21) and (33), allows one to conclude that the optimal value of the IDLP problem (28) and the optimal value of the dual problem (31) are equal (that is, the strong duality equality is valid) and also that $(\bar{\gamma},\bar{\xi})$ is an optimal solution of the former and $(\bar{\psi}(y), \bar{\eta}(y))$ is an optimal solution of the latter. Note that the common optimal value of problems (28) and (31) coincides with the Cesàro and Abel limits (27).

V. IMPROVEMENT OF THE ESTIMATES FROM BELOW

Proposition 5.1: The feasible set $\Omega(y_0)$ allows another (equivalent) representation in the form

$$\Omega(y_0) = \{(\gamma,\xi) \in \mathcal{P}(G) \times \mathcal{M}_+(G)|$$
$$\gamma_1(Q) - \int_G P(Q|y,u)\gamma(dy,du) = 0,$$
$$\xi_1(Q) - \int_G P(Q|y,u)\xi(dy,du) + \gamma_1(Q) = 1_Q(y_0)$$
$$\forall \text{ Borel } Q \subset Y\}, \tag{34}$$

where $P(Q|y,u)$ is the transition probability kernel corresponding to system (1), and where γ_1 and ξ_1 are marginals of γ and ξ, respectively.

Proof. The proof is given in [5]. $\square$

If the feasible set $\Omega(y_0)$ is presented in the form (34), then the dual problem to (17) takes the form (see [23]):

$$\sup_{(\mu,\psi,\eta)\in\hat{\mathcal{D}}(y_0)} \mu \stackrel{\text{def}}{=} \hat{d}^*(y_0), \tag{35}$$

where $\hat{\mathcal{D}}(y_0)$ is the set of triplets $(\mu, \psi(\cdot), \eta(\cdot)) \in I\!R \times \mathcal{B}(Y) \times \mathcal{B}(Y)$ that for all $(y,u) \in G$ satisfy the inequalities

$$k(y,u) +$$
$$(\psi(y_0) - \psi(y)) + E[\eta(f(y,u,s))] - \eta(y) - \mu \geq 0,$$
$$E[\psi(f(y,u,s))] - \psi(y) \geq 0,$$

where $\mathcal{B}(Y)$ stands for the space of bounded Borel functions on Y. Note that the following inequality holds

$$d^*(y_0) \leq \hat{d}^*(y_0) \leq k^*(y_0), \tag{36}$$

where the second inequality in (36) is just a version of the weak duality inequality, and it may be established similarly to Proposition 3.4.

The following proposition improves upon the estimates from below in Theorem 3.5.

Proposition 5.2: The following inequalities hold.

$$\liminf_{T\to\infty} v_T(y_0) \geq \hat{d}^*(y_0) \quad \forall\ y_0 \in Y,$$
$$\liminf_{\epsilon\to 0} h_\epsilon(y_0) \geq \hat{d}^*(y_0) \quad \forall\ y_0 \in Y.$$

Proof. The proof is given in [5]. $\square$

Corollary 5.3: Let, for a given $y_0 \in Y$, the strong duality equality be valid:

$$k^*(y_0) = \hat{d}^*(y_0). \tag{37}$$

Then the Cesàro and Abel limits of the optimal values exist and are equal:

$$\lim_{T\to\infty} v_T(y_0) = \lim_{\epsilon\to 0} h_\epsilon(y_0) = k^*(y_0) = \hat{d}^*(y_0). \tag{38}$$

Proof. Follows directly from Proposition 5.2 and Theorem 3.5. $\square$

Note that, if

$$k^*(y_0) = d^*(y_0) \tag{39}$$

(as in Example 1), then, by (36), $d^*(y_0) = \hat{d}^*(y_0) = k^*(y_0)$. That is, (37) is valid, with (38) taking a form identical to (23).

REMARK: Sufficient conditions for the equality (37) (the strong duality) to be valid have been studied in [23]. Note that the strong duality may not be true in the general case. An example, in which the Cesàro and Abel limits of the optimal values are not equal to each other, and, therefore, by Theorem 3.5, there is a duality gap, is given in [40].

VI. CONCLUSIONS

We have studied the problems of control of stochastic discrete-time systems with time averaging and time discounting optimality criteria. In particular, we have established that the Cesàro and Abel limits of the optimal values in such problems can be estimated with the help of a certain infinite-dimensional linear programming problem and its dual.

One possible future research direction is to find easily verifiable sufficient conditions for strong duality. Another interesting research direction is to relax some of the assumptions, e.g., the assumption about the compactness of the state space.

ACKNOWLEDGEMENT

We would like to thank Yi Zhang from the University of Liverpool for useful comments on the present work.

REFERENCES

[1] M. Arisawa and P.-L. Lions, On ergodic stochastic control, *Commun. in Partial Differential Equations*, 23:11 (1998), pp. 2187-2217.

[2] A. Arapostathis, V. S. Borkar and M. K. Ghosh, *Ergodic Control of Diffusion Processes*, Cambridge Uni. Press, Cambridge, UK, 2012.

[3] K. Avrachenkov, R.S. Burachik, J.A. Filar and V. Gaitsgory, "Constraint augmentation in pseudo-singularly perturbed linear programs", *Mathematical Programming, Ser. A*, 132 (2012), pp. 179-208.

[4] K. Avrachenkov, J.A. Filar, V. Gaitsgory and A. Stillman, "Singularly perturbed linear programs and Markov decision processes", *Operations Research Letters*, 44 (2016), pp. 297-301.

[5] K. Avrachenkov, V. Gaitsgory and L. Gamertsfelder, "LP Based Upper and Lower Bounds for Cesaro and Abel Limits of the Optimal Values in Problems of Control of Stochastic Discrete Time Systems", *Archived*, https://arxiv.org/abs/2010.15375v2,

[6] M. Bardi, I. Capuzzo-Dolcetta I *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1997.

[7] N. Bäuerle and U. Rieder, *Markov Decision Processes with Applications to Finance*, Springer, Berlin, 2011.

[8] C.J. Bishop, E.A. Feinberg and J. Zhang, Examples concerning Abel and Cesaro limits, *Journal of Mathematical Analysis and Applications*, 420 (2014), 1654-1661.

[9] A.G. Bhatt and V.S. Borkar, "Occupation Measures for Controlled Markov Processes: Characterization and Optimality", *Annals of Probability* 24 (1996), 1531-1562.

[10] V.S. Borkar, "A convex analytic approach to Markov Decision Processes", *Probability Theory and Related Fields* 78 (1988), 583-602.

[11] V. S. Borkar and V. Gaitsgory, Linear programming formulation of long run average optimal control problem, *Journal of Optimization Theory and Applications*, 181 (2019), pp. 101-125.

[12] V. S. Borkar, V. Gaitsgory and I. Shvartsman, LP Formulations of Discrete Time Long-Run Average Optimal Control Problems: the Nonergodic Case, *SIAM J. on Control and Otimization*, 57:3 (2019), 1783-1817.

[13] R. Buckdahn, D. Goreac, M. Quincampoix, Stochastic optimal control and linear programming approach. *Appl. Math. Optim.* 63 (2011), no. 2, 257–276.

[14] R. Buckdahn, M. Quincampoix and J. Renault, On representation formulas for long run averaging optimal control problem, *Journal of Differential Equations*, 259:11 (2015), pp. 5554-5581.

[15] W. H. Fleming and D. Vermes, "Convex duality approach to the optimal control of diffusions", *SIAM J. Control Optimization*, 27 (1989):5, pp. 1136-1155.

[16] V. Gaitsgory, "On Representation of the Limit Occupational Measures Set of a Control Systems with Applications to Singularly Perturbed Control Systems", *SIAM J. Control and Optimization*, 43 (2004):1, pp. 325-340.

[17] V. Gaitsgory, A. Parkinson and I.Shvartsman, "Linear Programming Formulations of Deterministic Infinite Horizon Optimal Control Problems in Discrete Time, *Discrete and Continuous Dynamical Systems, Series B*, 22 (2017):10, pp. 3821-3838.

[18] V. Gaitsgory, M. Quincampoix, "Linear programming approach to deterministic infinite horizon optimal control problems with discounting", *SIAM J. Control and Optim.*, vol. 48 (2009):4, pp 2480-2512.

[19] V. Gaitsgory, M. Quincampoix, "On sets of occupational measures generated by a deterministic control system on an infinite time horizon", *Nonlinear Analysis (Theory, Methods & Applications)*, 88 (2013), pp. 27-41.

[20] D. Goreac and O.-S.Serea, "Linearization techniques for L^∞ - control problems and dynamic programming principles in classical and L^∞ control problems", ESAIM: Control, Optimization and Calculus of Variations, 18 (2012):3, pp. 836-855.

[21] L. Grüne, Asymptotic controllability and exponential stabilization of nonlinear control systems at singular points, *SIAM J. Control Optim.*, 36(5), 1998, pp 1495-1503.

[22] L. Grüne, On the relation between discounted and average optimal value functions, *J. Diff. Equations*, 148, 1998, pp 65-69.

[23] O. Hernandez-Lerma and J. Gonzalez-Hernandez, "Infinite Linear Programming and Multichain Markov Control Processes in Uncountable Spaces", SIAM J. Control Optim., 36:1 (1998), pp. 313-335.

[24] D. Hernandez-Hernandez, O. Hernandez-Lerma, M. Taksar, "The Linear Programming Approach to Deterministic Optimal Control Problems", Applicationes Mathematicae, 24 (1996), No 1, pp 17-33.

[25] O. Hernandez-Lerma and J.B. Lasserre, "Linear Programming and Average Optimality of Markov Control Processes on Borel Spaces - Unbounded Costs", SIAM J. Control Optim., 32:2 (1994), pp. 480-500.

[26] O. Hernandez-Lerma and J.B. Lasserre, "Discrete-Time Markov Control Processes", Springer, New York, 1996.

[27] O. Hernandez-Lerma and J.B. Lasserre, "Further Topics on Discrete-Time Markov Control Processes", Springer, New York, 1999.

[28] A Hordijk and L. C. M. Kallenberg, Linear programming and Markov decision chains, *Manage. Sci.*, 25:4 (1979), pp 352-362.

[29] A. Hordijk and L. C. M. Kallenberg, Constrained undiscounted stochastic dynamic programming, *Math. Oper. Res.* 9:2 (1984), 276-289.

[30] D. Khlopin, Tauberian theorem for value functions, *Dyn. Games Appl.*, 8 (2018), pp. 401-422.

[31] Kurtz T.G. and Stockbridge R.H., "Existence of Markov Controls and Characterization of Optimal Markov Controls", *SIAM J. on Control and Optimization*, 36:2 (1998), pp. 609-653.

[32] J.B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, Nonlinear optimal control via occupation measures and LMI-relaxations, SIAM J. Control Optim., 47 (2008), pp. 1643-1666.

[33] E. Lehrer, S. Sorin, A uniform Tauberian theorem in dynamic programming, *Mathematics of Operations Research*, 17(2), (1992), 303-307.

[34] M. Oliu-Barton and G. Vigeral, A uniform Tauberian theorem in optimal control, In P. Cardaliaguet and R. Grossman (Eds.)*Annals of International Society of Dynamic Games*, 12, pp. 199-215, Birkhauser/Springer, New York, 2013.

[35] A.B. Piunovskiy, *Optimal Control of Random Sequences in Problems with Constraints*, Springer, 1997.

[36] A. Piunovskiy and Y. Zhang, Discounted Continuous-Time Markov Decision Processes with Unbounded Rates: The Convex Analytic Approach, *SIAM J Control Optim.* 49:5 (2011), 2032-2061.

[37] M. Quincampoix and J. Renault, On existence of a limit value in some non-expansive optimal control problems, *SIAM J. on Control and Optimization*, 49:5 (2012), pp. 2118-2132.

[38] J.E. Rubio, "Control and Optimization. The Linear Treatment of Nonlinear Problems", Manchester University Press, Manchester, 1985.

[39] R.H. Stockbridge, "Time-Average Control of a Martingale Problem: A Linear Programming Formulation, *Annals of Probability*, 18.

[40] I. Shvartsman, Lack of Equality between Abel and Cesaro Limits in Discrete Optimal Control and the Implied Duality Gap, *Archived*, http://arxiv.org/abs/2010.14676

[41] R. Sznajder, J.A. Filar, *Some comments on a theorem of Hardy and Littlewood*, J. Optimization Theory and Applications, 75(1), (1992) 201-208.

[42] R. Vinter, "Convex Duality and Nonlinear Optimal Control", *SIAM J. Control and Optim.* Vol. 31 (1993), No 2, pp. 518-538.