A Convex Duality Approach for Assigning Probability Distributions to the State of Nonlinear Stochastic Systems

Ali Pakniyat, *Member, IEEE*

***Abstract*—In order to optimally assign a desired probability distribution to the state of a nonlinear stochastic system, a convex duality approach is proposed to arrive at the associated optimality conditions. For a general class of stochastic systems governed by controlled Itô differential equations and subject to constraints on the probability distribution of the state at a fixed terminal time, a measure theoretic formulation is presented and it is shown that the original problem is embedded in a convex linear program on the space of Radon measures and that the embedding is tight, i.e., the optimal solution of both the original and the convex relaxation problems are equal. By exploiting the duality relationship between the space of continuous functions and that of measures, the associated optimality conditions are identified in the form of Hamilton-Jacobi problems where the optimization objective, in addition to the value function evaluation at the initial conditions, includes an extra term which is the integral of the product of the value function at the terminal time and the desired probability distribution. Numerical examples are provided to illustrate the results.**

***Index Terms*—Stochastic optimal control, stochastic systems, optimal control.**

I. Introduction

THIS letter addresses finite-horizon optimal control problems for continuous time nonlinear stochastic systems, where the control objective is to steer the state from an initial condition to a desired terminal probability distribution with known statistics. In the literature, problems of this type has only appeared for special subclasses of systems. More precisely, the majority of studies assume linearity of the dynamics and a quadratic form for the cost, as well as Gaussian forms for the desired distribution. The associated results are presented for both infinite time horizon problems [1]–[4] and finite time horizon problems in both continuous time and discrete time settings [5]–[13]. The accommodation of input constraints is considered in [9], and convex relaxations for linear systems subject to chance constraints are studied in [11], [12]. Extensions of the probability distribution assignment to nonlinear systems has been presented for feedback-linearizable systems [14], and implementation through iterative linearization is proposed in [15]. In past work of the author [16], [17], a perpetual renewal of the terminal state constraint is proposed for nonlinear stochastic systems and the Terminally Constrained Stochastic Minimum Principle (TC-SMP) is established. While the TC-SMP applies to both linear and nonlinear systems, a major limitation of the theory is that it can only assigns the Dirac delta probability distribution to the terminal state and this is achieved at the expense of using unbounded input values. In this letter, we take an entirely different approach which permits assigning a large class of probability distributions to the terminal state. We employ the notion of occupation measures to formulate the problem as a linear program over the space of measures and we arrive at the optimality conditions by using convex duality relations.

The convex duality method for optimal control problems was initiated by Vinter and Lewis [18], [19] for deterministic control systems and, later, by Fleming and Vermes for piecewise deterministic [20] and stochastic [21] processes. The fundamental idea of this approach is the introduction of a weak formulation that embeds the original (strong) problem into a convex linear program over the space of Radon measures. Upon establishing the equivalence of the two problems, new necessary and sufficient optimality condition are obtained by invoking the Fenchel-Rockafellar duality theorem. This approach is particularly useful in characterization of optimal policies in certain desirable classes of controls by investigating the extreme points of the set of Hamilton-Jacobi problems (see, e.g., [22]–[25]). For deterministic control systems, convex duality based numerical algorithms are established in [26]–[28] for continuous systems, and in [29]–[31] for hybrid systems.

The first objective of this letter is the extension of the covariance control problems to nonlinear systems with nonlinear costs and general desired probability distributions. This, in particular, requires a change of viewpoint from the study of *sample paths* (where the terminal state distribution is a statistical byproduct of the investigation) to the study of the so-called

Manuscript received March 22, 2022; revised May 12, 2022; accepted May 30, 2022. Date of publication June 8, 2022; date of current version June 22, 2022. Recommended by Senior Editor V. Ugrinovskii.

The author is with the Department of Mechanical Engineering, University of Alabama, Tuscaloosa, AL 35401 USA (e-mail: apakniyat@ua.edu).

Digital Object Identifier 10.1109/LCSYS.2022.3181525

occupation measures in which the description of the terminal distribution constraint is inherently natural to the representation. The second and primary objective of the article is to establish the associated optimality conditions for the general nonlinear and non-Gaussian case which itself requires invoking convex duality relationships. While the majority of this part has parallels in the discussion for unconstrained nonlinear stochastic systems in [21], our optimality conditions for nonlinear stochastic systems with distribution constraints are novel and, to the best of our knowledge, has not appeared in the literature before. The third objective is to illustrate the theoretical results by means of numerical examples.

The organization of this letter is as follows. Section II presents the class of stochastic optimal control problems subject to distribution constraints on their terminal state. Section III introduces the notion of occupation measures and shows that the cost can be strongly represented as a linear functional over the space of occupation measures. Since the identification of this space is not straightforward from the governing equations, Section IV investigates the properties of occupation measures based upon which a weak formulation of the problem is presented in Section V as a linear program (LP) defined on a convex domain in the space of signed measures. It is shown that the strong problem is tightly embedded in this weak problem and a set of Hamilton Jacobi (HJ)-type inequalities are established using the duality relationship between the space of measures and that of continuous functions. Illustrative examples are provided in Section VI.

II. Original Problem

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_0^T, \mathsf{P})$ be a filtered probability space with $\{\mathcal{F}_t\}_0^T$ an increasing family of sub σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_0$ contains all the P-null sets, and $\mathcal{F}_T = \mathcal{F}$ for the fixed terminal time $T < \infty$. Consider a nonlinear stochastic systems governed by the controlled Itô differential equation

$$\mathsf{d}x_s = f(s, x_s, u_s)\mathsf{d}s + g(s, x_s)\mathsf{d}w_s, \tag{1}$$

where $x_s \in \mathbb{R}^n$, $u_s \in U \subset \mathbb{R}^m$, and $w_s \in \mathbb{R}^d$ are, respectively, the values of the state, the input, and the realization of a standard Wiener process at time $s \in [0, T]$. The input value set U is assumed to be convex and compact and the functions f and g are considered to be Lipschitz functions over, respectively, $[0, T] \times \mathbb{R}^n \times U$ and $[0, T] \times \mathbb{R}^n$ with linearly bounded growth rates.

Let $[u] := \{u_s : 0 \leq s \leq T\}$ denote a nonanticipative, U-valued, input process such that $u_s \in U$ is progressively measurable with respect to $\mathcal{F}_s$ for all $s \in [0, T]$. We denote by $\mathcal{U}$ the set of all such inputs.

The primary goal of the control problem is to determine $[u] \in \mathcal{U}$ such that the probability distribution of the terminal state takes a desired form p_d, i.e., $x_T^{[u]} \sim \mathsf{p}_d$. This, by definition, signifies that for every Borel set $B_x \in \mathbb{R}^n$,

$$\mathsf{P}^{[u]}(x_T \in B_x) = \int_{B_x} \mathsf{p}_d(\mathsf{d}x), \tag{2}$$

where $\mathsf{P}^{[u]}(\cdot)$ denotes the probability of an event given the input $[u]$. The set of all $[u] \in \mathcal{U}$ such that (2) is satisfied is denoted by $\mathcal{U}'$. In this letter, we assume that the system satisfies any controllability requirement so that the desired probability distribution is attainable. More precisely, we restrict our attention to problems satisfying the following.

Assumption 1: The set $\mathcal{U}'$ is non-empty.

For a given initial condition x_0 at $t = 0$, we associated to each $[u] \in \mathcal{U}'$ a total cost

$$J(0, x_0, [u]) = \mathbb{E}^{[u]}\left[\int_0^T \ell(x_s, u_s)\mathsf{d}s\right] \tag{3}$$

where ℓ is a continuous function with polynomial growth.

The associated optimal control problem is defined as finding the value function at the initial time and state

$$V(0, x_0) := \inf_{[u]}\left\{\mathbb{E}\left[\int_0^T \ell(x_s, u_s)\mathsf{d}s \,\middle|\, [u]\right] \text{ s.t. } x_T^{[u]} \sim \mathsf{p}_d\right\} \tag{P}$$

and, whenever a minimizer exists, an optimal policy $[u^o]$ which attains the minimum value of (P).

III. Strong Formulation

We define the *input-state-time occupation measure* as

$$\mu^{[u]}(B_t, B_x, B_u) := \mathbb{E}^{[u]} \int_{B_t} \mathbb{I}_{B_x}(x_s) \cdot \mathbb{I}_{B_u}(u_s)\, \mathsf{d}s, \tag{4}$$

for arbitrary Borel sets $B_t \subset [0, T]$, $B_x \subset \mathbb{R}^n$, $B_u \subset U$, where $\mathbb{I}_B$ denotes the indicator function of the set B.

We also define the *terminal state occupation measure* as

$$\kappa^{[u]}(B_x) := \mathsf{P}^{[u]}(x_T \in B_x). \tag{5}$$

for an arbitrary Borel set $B_x \subset \mathbb{R}^n$.

Lemma 1: For every $[u] \in \mathcal{U}$, measurable functions $\ell : [0, T) \times \mathbb{R}^n \times U \to \mathbb{R}$ with $\ell(s, x, U) := \{l(s, x, u) : u \in U\}$ convex for all $s \in [0, T]$, $x \in \mathbb{R}^n$, and for all measurable functions $L : \mathbb{R}^n \to \mathbb{R}$, it is the case that

$$\mathbb{E}^{[u]} \int_0^T \ell(x_s, u_s)\mathsf{d}s = \int_{[0,T]\times\mathbb{R}^n\times U} \ell(x, u)\, \mu^{[u]}(\mathsf{d}t, \mathsf{d}x, \mathsf{d}u) =: \left\langle \ell, \mu^{[u]}\right\rangle, \tag{6}$$

$$\mathbb{E}^{[u]}[L(x_T)] = \int_{\mathbb{R}^n} L(x)\, \kappa^{[u]}(\mathsf{d}x) =: \left\langle L, \kappa^{[u]}\right\rangle. \tag{7}$$

Proof: The proof follows directly from the definitions (4) and (5), and the measurability of the functions ℓ, L. ■

We denote by $\mathcal{M}_S$ the set of occupations measures corresponding to all $[u] \in \mathcal{U}$, i.e., $\mathcal{M}_S := \{\mu^{[u]} : [u] \in \mathcal{U}\}$.

Thus, the original problem (P) is equivalently represented in terms of occupation measures in the form of

$$V(0, x_0) = \inf_{\mu^{[u]} \in \mathcal{M}_S} \left\{\left\langle \ell, \mu^{[u]}\right\rangle \text{ s.t. } \kappa^{[u]} = \mathsf{p}_d\right\}. \tag{SP}$$

We refer to the reformulation (SP) as the *strong problem* due to the direct correspondence between (P) and (SP). We note that for every measurable function ℓ, the problem (SP) is an optimization problem with a linear objective defined over the space $\mathcal{M}_S$. However, the identification of this space is not straightforward as it is associated with implementing all admissible inputs $[u] \in \mathcal{U}$ on the stochastic differential equation (1). To address this issue, we present in Section V a problem defined directly over the space of measures which tightly embeds our original problem.

IV. From Strong to Weak Formulation

The goal of this section is to invoke the properties of occupation measures in Section III to provide a formulation of the strong problem (SP) as a linear program over a convex subspace in the space of measures presented in Section V.

A. Positivity of Occupation Measures

The first observation about the occupation measures (4) is that they are non-negative measures. In other words, for every $[u] \in \mathcal{U}$ it is the case that $\mu^{[u]}(B_t, B_x, B_u) \geq 0$, for all $B_t \subset [0, T]$, $B_x \subset \mathbb{R}^n$, $B_u \subset U$.

B. Finiteness of Measure Norms

The norm of occupation measures defined over a finite time horizon is bounded by

$$\begin{aligned} \|\mu^{[u]}\| &= \int_{[0,T]\times\mathbb{R}^n\times U} \mu^{[u]}(\mathsf{d}s, \mathsf{d}x, \mathsf{d}u) \\ &\underset{(4)}{=} \int_0^T \int_{\mathbb{R}^n} \int_U \mathbb{I}_{\mathsf{d}x}(x_s) \cdot \mathbb{I}_{\mathsf{d}u}(u_s)\, \mathsf{d}s = T, \end{aligned} \tag{8}$$

C. Governing Equations

Before presenting the equations governing the evolution of occupation measures, let us first present the following.

Lemma 2: For every twice continuously differentiable function $v \in C^2([0, T] \times \mathbb{R}^n)$

$$\langle v, \kappa^{[u]} \rangle - \langle \mathcal{A}v, \mu^{[u]} \rangle = v(0, x_0), \tag{9}$$

where $\mathcal{A}$ is the infinitesimal operator of the Markov process (1) and is given as

$$\begin{aligned} \mathcal{A}^u v(t, x) &= \frac{\partial v(t, x)}{\partial t} + \left[\frac{\partial v(t, x)}{\partial x}\right]^\top f(t, x, u) \\ &\quad + \frac{1}{2}\mathrm{tr}\left(g(t, x)^\top g(t, x)\, \frac{\partial^2 v(t, x)}{\partial x^2}\right). \end{aligned} \tag{10}$$

Proof: It follows from Dynkin's formula that (see, e.g., [21, eq. (8.3)])

$$\mathbb{E}^{[u]}[v(T, x_T)] = v(0, x_0) + \mathbb{E}^{[u]}\left[\int_0^T \mathcal{A}^{u_s} v(s, x_s)\, \mathsf{d}s\right] \tag{11}$$

for all $[u] \in \mathcal{U}$. Invoking Lemma 1, we can rewrite the left and right hand side terms in (11) as

$$\mathbb{E}^{[u]}[v(T, x_T)] = \int_{\mathbb{R}^n} v(T, x)\, \kappa^{[u]}(\mathsf{d}x) = \langle v, \kappa^{[u]} \rangle \tag{12}$$

$$\begin{aligned} &\mathbb{E}^{[u]}\left[\int_0^T \mathcal{A}^{u_s} v(s, x_s)\, \mathsf{d}s\right] \\ &= \int_{[0,T]\times\mathbb{R}^n\times U} \mathcal{A}^u v(t, x)\, \mu^{[u]}(\mathsf{d}t, \mathsf{d}x, \mathsf{d}u) = \langle \mathcal{A}v, \mu^{[u]} \rangle. \end{aligned} \tag{13}$$

Substitution of (12) and (13) into (11) yields (9). ■

Using Lemma 2, we can write the equation governing the evolution of occupation measures as follows.

Theorem 1: For every $[u] \in \mathcal{U}$, occupation measures corresponding to trajectories of the system (1) satisfy

$$\kappa^{[u]} - \mathcal{A}^* \mu^{[u]} = \bar{\delta}_{(0,x_0)}, \tag{14}$$

where $\bar{\delta}_{(0,x_0)}$ is the Dirac measure, and $\mathcal{A}^*$ is the adjoint of (10) defined as the operator satisfying

$$\langle \mathcal{A}v, \mu \rangle = \langle v, \mathcal{A}^* \mu \rangle \tag{15}$$

for every Borel measure μ, and any twice continuously differentiable function $v \in C^2([0, T) \times \mathbb{R}^n)$.

Proof: By invoking Lemma 2 and the definition of the Dirac measure, (9) is written as

$$\langle v, \kappa^{[u]} \rangle - \langle \mathcal{A}v, \mu^{[u]} \rangle = \langle v, \bar{\delta}_{(0,x_0)} \rangle. \tag{16}$$

Since all $\mu^{[u]} \in \mathcal{M}_S$ are Borel measures, we can invoke (15) for every $\mu^{[u]}$ and obtain

$$\langle v, \kappa^{[u]} \rangle - \langle v, \mathcal{A}^* \mu^{[u]} \rangle = \langle v, \bar{\delta}_{(0,x_0)} \rangle, \tag{17}$$

which, by the additive property of inner products, becomes

$$\langle v, \kappa^{[u]} - \mathcal{A}^* \mu^{[u]} - \bar{\delta}_{(0,x_0)} \rangle = 0. \tag{18}$$

Since (18) must hold for all $v \in C^2([0, T) \times \mathbb{R}^n)$, the relation (14) is obtained. ■

In particular, the substitution of $\kappa^{[u]} = \mathsf{p}_d$ into (14) yields

$$\mathcal{A}^* \mu^{[u]} = \mathsf{p}_d - \bar{\delta}_{(0,x_0)}. \tag{19}$$

V. Weak Formulation

Since the identification of $\mathcal{M}_S$ is challenging due to its dependence on $\mathcal{U}$, in this section, we introduce a weaker problem define over a larger domain $\mathcal{M}_W \supset \mathcal{M}_S$ that is easily identifiable as a convex domain in the space of measures.

A. Weak Problem

Let $\mathfrak{M}_\pm(S)$ denote the set of all signed Borel measures on S and $\mathfrak{M}_+(S)$ the non-negative cone of $\mathfrak{M}_\pm(S)$. For every $\mu \in \mathfrak{M}_\pm([0, T] \times \mathbb{R}^n \times U)$ we define the norm $\|\mu\| := \int_{[0,T]\times\mathbb{R}^n\times U} \mathsf{d}\mu^+ + \int_{[0,T]\times\mathbb{R}^n\times U} \mathsf{d}\mu^-$.

We begin by defining the weak problem and the corresponding weak value function as

$$W(0, x_0) := \inf_{\mu \in \mathcal{M}_W} \langle \ell, \mu \rangle, \tag{20}$$

where $\mathcal{M}_W := \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}}$, with

$$\mathcal{M}_{PB} := \{\mu \in \mathfrak{M}_+([0, T] \times \mathbb{R}^n \times U) : \|\mu\| \leq T\}, \tag{21}$$

$$\mathcal{M}_{\mathcal{A}} := \{\mu \in \mathfrak{M}_\pm([0, T] \times \mathbb{R}^n \times U) : \mathcal{A}^*\mu = \mathsf{p}_d - \bar{\delta}_{(0,x_0)}\}. \tag{22}$$

The above problem is a linear program on the space of signed measures. The set $\mathcal{M}_{PB}$ is a convex subset of $\mathfrak{M}_\pm$ and the constraint $\mathcal{M}_{\mathcal{A}}$ is linear and therefore restricts the problem into a linear subspace.

Over the compact Hausdorff space $[0, T] \times \mathbb{R}^n \times U$, the Banach space of continuous functions $C([0, T] \times \mathbb{R}^n \times U)$ equipped with the sup-norm has a topological dual $C^*([0, T] \times \mathbb{R}^n \times U)$ that is isometrically isomorphic to $\mathfrak{M}_\pm([0, T] \times \mathbb{R}^n \times U)$ equipped with the norm $\|\mu\| := \int \mathsf{d}\mu^+ + \int \mathsf{d}\mu^-$. The norm topology of C and the weak dual topology of $\mathfrak{M}_\pm$ are compatible with the pairing defined by the bilinear form $\langle c, \mu \rangle$ for all $c \in C([0, T] \times \mathbb{R}^n \times U)$, and $\mu \in \mathfrak{M}_\pm([0, T] \times \mathbb{R}^n \times U)$.

Endowing the space of continuous functions with the topology of the sup-norm and endowing the space of signed measures, $\mathfrak{M}_\pm$, with a weak dual topology, it follows that $\mathcal{M}_{PB} \cap \mathcal{M}_\mathcal{A}$ is w*-compact and hence, the infimum in (20) is achieved and is equal to the minimum. Thus, we define the weak problem as

$$W(0,x_0) := \min_{\mu\in\mathfrak{M}_\pm} \{\langle \ell,\mu\rangle, \text{ s.t. } \mu \in \mathcal{M}_{PB}\cap\mathcal{M}_\mathcal{A}\}. \quad \text{(WP)}$$

B. Fenchel Normal Form

Using the notion of weak value function, we reformulate the convexly constrained linear program as an unconstrained convex problem by introducing the functionals h_1 and $h_2 : \mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U) \to \overline{\mathbb{R}}$ defined by

$$h_1(\mu) := \begin{cases} \langle \ell,\mu\rangle, & \text{if } \mu\in\mathcal{M}_{PB}, \\ +\infty, & \text{otherwise.}\end{cases} \tag{23}$$

$$h_2(\mu) := \begin{cases} 0, & \text{if } \mu\in\mathcal{M}_{\mathcal{A}}, \\ -\infty, & \text{otherwise.}\end{cases} \tag{24}$$

Both h_1 and $-h_2$ are convex and lower semi-continuous [21] and, hence,

$$W(0,x_0) = \min_{\mu\in\mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U)} \{h_1(\mu) - h_2(\mu)\}. \tag{25}$$

C. Legendre-Fenchel Transform

The real-valued functional h_1 is convex and its convex conjugate (Legendre-Fenchel transform) is defined by

$$h_1^*(c) := \sup_{\mu\in\mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U)} \{\langle c,\mu\rangle - h_1(\mu)\}. \tag{26}$$

Lemma 3:

$$h_1^*(c) = T\cdot \|(c-\ell)^+\|, \tag{27}$$

where $(f)^+$ denotes the positive part of the function f, i.e., $f^+(x) = max\{0, f(x)\}$.

Proof: The proof is a modification of [21, Lemma 4.1] and is removed due to space limitations. ■

For the concave functional h_2 the Legendre-Fenchel transform is defined as

$$h_2^*(c) := \inf_{\mu\in\mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U)} \{\langle c,\mu\rangle - h_2(\mu)\}. \tag{28}$$

Lemma 4:

$$h_2^*(c) = \begin{cases} \lim_{i\to\infty}\left(v_i(0,x_0) - \langle v_i, \mathsf{p}_d\rangle\right), & \text{if } c = -\lim_{i\to\infty}\mathcal{A}v_i, \\ -\infty, & \text{otherwise.}\end{cases} \tag{29}$$

Proof: The proof is a modification of [21, Lemma 4.2] and is removed due to space limitations. ■

D. The Hamilton-Jacobi Problem

Theorem 2:

$$W(0,x_0) = \sup_{v\in C^2([0,T]\times\mathbb{R}^n)} \Big\{ v(0,x_0) - \int_{\mathbb{R}^n} v(T,x)\mathsf{p}_d(\mathrm{d}x), \quad \text{s.t. } \mathcal{A}v + \ell \geq 0 \Big\}. \tag{30}$$

Proof: Applying the Rockafellar duality theorem [21] to $C^*([0,T]\times\mathbb{R}^n\times U) = \mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U)$, we obtain

$$\min_{\mu\in\mathfrak{M}_\pm([0,T]\times\mathbb{R}^n\times U)} \{h_1(\mu) - h_2(\mu)\}$$
$$= \sup_{c\in C([0,T]\times\mathbb{R}^n\times U)} \{h_2^*(c) - h_1^*(c)\} \tag{31}$$

whenever the set $\{c : h_2^*(c) > -\infty\}$ contains a continuity point of $h_1^*(c)$ that is finite. Since h_1^* is continuous and finite on whole $C([0,T]\times\mathbb{R}^n\times U)$ and h_2^* is not identically $-\infty$ we deduce that (31) holds. The substitution of (31) into (25) yields

$$\begin{aligned} W(0,x_0) &= \sup_{c\in C([0,T]\times\mathbb{R}^n\times U)} \{h_2^*(c) - h_1^*(c)\} \\ &\overset{(27)}{\underset{(29)}{=}} \sup_{c\in C} \Big\{ \lim_{i\to\infty}\left(v_i(0,x_0) - \langle v_i,\mathsf{p}_d\rangle\right) - T\|(c-\ell)^+\| \\ &\qquad \text{s.t. } c = -\lim_{i\to\infty}\mathcal{A}v_i \Big\}. \end{aligned} \tag{32}$$

Using the fact that $\{\mathcal{A}v : v\in C^2([0,T]\times\mathbb{R}^n)\}$ is dense in $\{c\in C([0,T]\times\mathbb{R}^n\times U) : h_2^*(c) > -\infty\}$, we obtain

$$W(0,x_0) = \sup_{v\in C^2} \Big\{ \left(v(0,x_0) - \langle v,\mathsf{p}_d\rangle\right) - T\|(c-\ell)^+\| \quad \text{s.t. } c = -\mathcal{A}v \Big\}. \tag{33}$$

To conclude the proof it suffices to show that for every $v \in C^2([0,T]\times\mathbb{R}^n)$ there exists a $\tilde{v} \in C^2([0,T]\times\mathbb{R}^n)$ such that $\mathcal{A}\tilde{v} + \ell \geq 0$ and $\tilde{v}(0,x_0) \geq v(0,x_0)$. Defining $\tilde{v} := v - T\|(\mathcal{A}v+\ell)^-\|$, it follows that

$$\mathcal{A}\tilde{v} + \ell \equiv \mathcal{A}v + \ell + \|(\mathcal{A}v+\ell)^-\| \geq \mathcal{A}v + \ell + \sup_{(s,x,u)\in[0,T]\times X\times U} |(\mathcal{A}^u v(s,x) + \ell(x,u))^-| \geq 0. \tag{34}$$

■

E. Equivalence of the Weak and Strong Problems

It follows from the definitions (WP) and (SP) of the weak and strong value functions that

$$W(0,x_0) = \min_{\mu\in\mathcal{M}_W} \langle \ell,\mu\rangle \leq V(0,x_0) = \inf_{\mu^{[u]}\in\mathcal{M}_S} \{\langle \ell,\mu^{[u]}\rangle\}. \tag{35}$$

since $\mathcal{M}_S \subset \mathcal{M}_W := \mathcal{M}_{PB}\cap\mathcal{M}_\mathcal{A}$. In order to show the equivalence of the weak and the strong problems, we need to show that strict inequality cannot hold and hence, the weak and the strong value functions coincide.

Theorem 3: The weak and the strong value functions are equal, i.e.,

$$W(0,x_0) = V(0,x_0). \tag{36}$$

Proof: Let's assume that this is not true, i.e., there exist $(\mu_0,\kappa_0) \in \mathcal{M}_{PB}\cap\mathcal{M}_\mathcal{A}\backslash\mathcal{M}_S$ such that

$$W_0(0,x_0) = \langle \ell,\mu_0\rangle < V(0,x_0) = \inf_{\mu^{[u]}\in\mathcal{M}_S} \{\langle \ell,\mu^{[u]}\rangle\}. \tag{37}$$

This means that the w*-continuous linear functional $\langle \ell,\mu\rangle$ separates an element $\mu_0 \in \mathcal{M}_{PB}\cap\mathcal{M}_\mathcal{A}$ from the w* convex closure $\overline{\text{cov}\mathcal{M}_S}$ of $\mathcal{M}_S$. Then by [21, Th. 3], for every $\epsilon > 0$ there exists $V^{(\epsilon)}$ whose partial derivatives $V_t^{(\epsilon)}$, $V_{x_i}^{(\epsilon)}$, $V_{x_ix_j}^{(\epsilon)}$ are defined almost everywhere, are essentially bounded and, further,

$$\|V - V^{(\epsilon)}\| \leq \epsilon,\ \mathcal{A}^u V^{(\epsilon)}(s,x) + \ell(x,u) \geq 0, \tag{38}$$

for all $(s, x, u) \in [0, T] \times \mathbb{R}^n \times U$. Since $V^{(\epsilon)}$ is not necessarily in $C^2([0, T] \times \mathbb{R}^n \times U)$, in order to apply Dynkin's formula (9), we also need to invoke [21, Lemma 5.1] that for every $\delta > 0$ there exists $V^{(\epsilon,\delta)} \in C^2([0, T] \times \mathbb{R}^n \times U)$ such that

$$\|V^{(\epsilon,\delta)} - V^{(\epsilon)}\| < \delta, \|\mathcal{A}V^{(\epsilon,\delta)}\| \le \|\mathcal{A}V^{(\epsilon)}\| + \delta, \quad (39)$$
$$\mathcal{A}V^{(\epsilon,\delta)} + \ell \ge -\delta, \text{ on } [\delta, T-\delta] \times \mathbb{R}^n \times U. \quad (40)$$

Then by (9),

$$\begin{aligned} V^{(\epsilon,\delta)}(0, x_0) - \big\langle V_T^{(\epsilon,\delta)}, \mathsf{p}_d\big\rangle &= -\big\langle \mathcal{A}V^{(\epsilon,\delta)}, \mu_0\big\rangle \\ &\le \int_{[\delta,T-\delta)\times\mathbb{R}^n\times U} \ell \, \mathsf{d}\mu_0 + \delta \int_{[\delta,T-\delta)\times\mathbb{R}^n\times U} \mathsf{d}\mu_0 \\ &+ \|\mathcal{A}V^{(\epsilon,\delta)}\| \int_{\{[0,\delta)\cup[T-\delta,T)\}\times\mathbb{R}^n\times U} \mathsf{d}\mu_0 \end{aligned} \quad (41)$$

and hence,

$$V^{(\epsilon,\delta)}(0, x_0) - \big\langle V_T^{(\epsilon,\delta)}, \mathsf{p}_d\big\rangle \le \langle \ell, \mu_0\rangle + 2 \cdot \delta \cdot T\Big(1 + \|\mathcal{A}V^{(\epsilon,\delta)}\|\Big) \quad (42)$$

Employing $\|V - V^{(\epsilon,\delta)}\| < \epsilon + \delta$ from (38) and (39), and choosing first ϵ then δ sufficiently small, we arrive at

$$V(0, x_0) - \big\langle V_T^{(\epsilon,\delta)}, \mathsf{p}_d\big\rangle \le \langle \ell, \mu_0\rangle \quad (43)$$

that is in contradiction with the hypothesis (37). Therefore, the equivalence (36) holds true. ■

Theorem 4 (Main Result): For every $x_0 \in \mathbb{R}^n$ and given a desired terminal distribution p_d, the optimal cost (P) is obtained as

$$\begin{aligned} V(0, x_0) = \sup_{v\in C^2([0,T]\times\mathbb{R}^n)} \Bigg\{ & v(0, x_0) - \int_{\mathbb{R}^n} v(T, x)\mathsf{p}_d(\mathsf{d}x), \\ \text{s.t.} \quad & \frac{\partial v(t, x)}{\partial t} + \left[\frac{\partial v(t, x)}{\partial x}\right]^\top f(t, x, u) \\ & + \frac{1}{2}\mathrm{tr}\left(g(t, x)^\top g(t, x) \, \frac{\partial^2 v(t, x)}{\partial x^2}\right) + \ell(t, x, u) \ge 0, \\ & \text{for all } (t, x, u) \in [0, T] \times \mathbb{R}^n \times U \Bigg\}. \end{aligned} \quad (44)$$

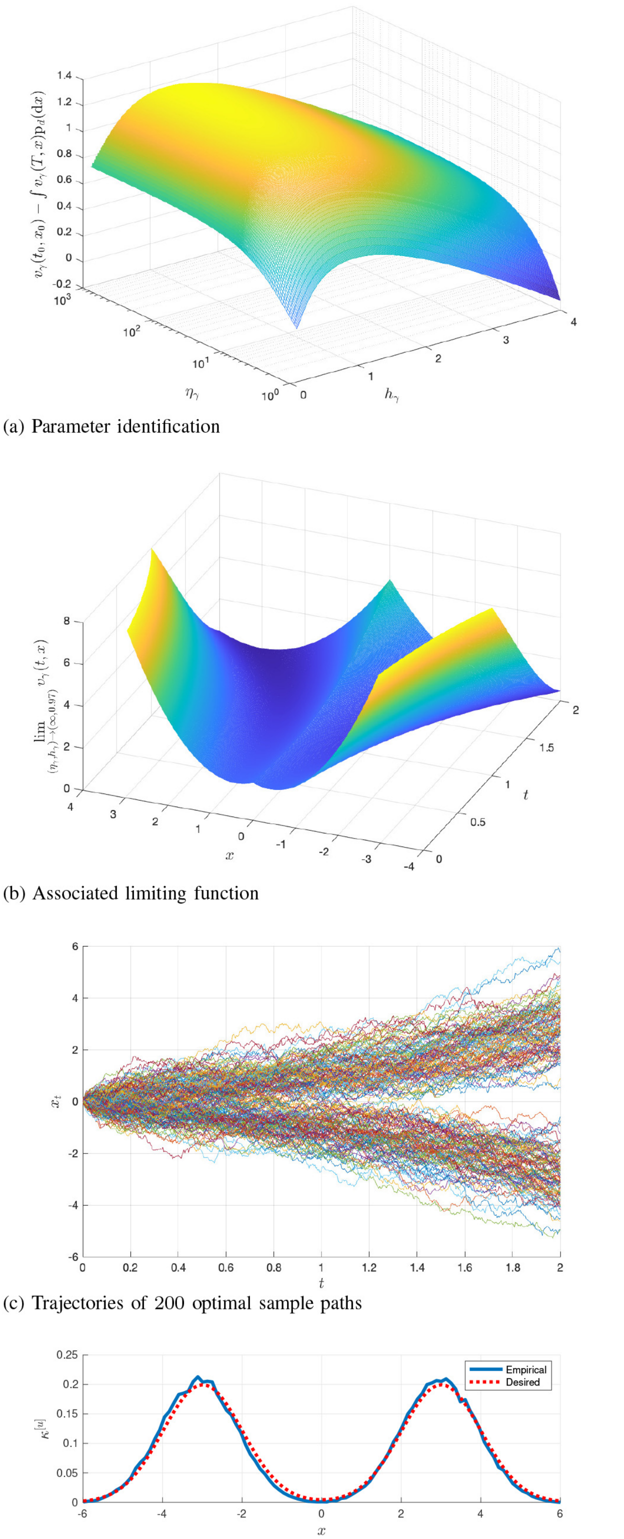
Proof: The result is obtained by substituting (36) from Theorem 3 into (30) from Theorem 2. ■

VI. Numerical Illustration

Consider the scalar system $\mathsf{d}x_s = (x_s + u_s)\mathsf{d}s + \mathsf{d}w_s$, with the total cost $J(t_0, x_0, [u]) = \mathbb{E}\int_0^T \frac{1}{2}u_s^2 \mathsf{d}s$ and the desired terminal distribution $\mathsf{p}_d = \frac{1}{2}\mathcal{N}(\mu_1^d, \sigma_1^d) + \frac{1}{2}\mathcal{N}(\mu_2^d, \sigma_2^d)$. These problems, despite their LQ form of the dynamics and cost, cannot be solved by the conventional covariance control methodologies [1]–[13]. In contrast, the results of Theorem 4 can be implemented in the following way to identify the value function and the corresponding optimal policy.

Consider the family of functions $\{v_\gamma\}$ where for each $\gamma = (\eta_\gamma, \rho_\gamma, h_1^\gamma, \mu_1^\gamma, h_2^\gamma, \mu_2^\gamma)$, the function is expressed as

$$\begin{aligned} v_\gamma(t, x) = \frac{-1}{\eta_\gamma} \ln\bigg(& e^{-\eta_\gamma\rho_\gamma\left(\frac{1}{2}\pi_1(t)x^2 + \beta_1(t)x + \alpha_1(t)\right)} \\ & + e^{-\eta_\gamma(1-\rho_\gamma)\left(\frac{1}{2}\pi_2(t)x^2 + \beta_2(t)x + \alpha_2(t)\right)}\bigg) \end{aligned} \quad (45)$$



(a) Parameter identification

(b) Associated limiting function

(c) Trajectories of 200 optimal sample paths

(d) Distribution of x_T for 100‘000 sample paths vs. desired distribution

Fig. 1. The identification of parameters (a), the associated limiting function (b), sample paths (c), and the corresponding distribution (d) for $x_0 = 0$, $T = 2$ and $\mathsf{p}_d = \frac{1}{2}\mathcal{N}(3,\ 1) + \frac{1}{2}\mathcal{N}(-3,\ 1)$ employing (44) and the class of functions $\{v_\gamma\}$ defined by (45).

where π_i, β_i, α_i, $i = 1, 2$, satisfy the Riccati equations $\dot{\pi}_i = \pi_i^2 - 2\pi_i$, $\pi_i(T) = h_i^\gamma$; $\dot{\beta}_i = -(1 - \pi_i)\beta_i$, $\beta_i(T) = -h_i^\gamma\mu_i^\gamma$; $\dot{\alpha}_i = \frac{1}{2}\beta_i^2 - \frac{1}{2}\pi_i$, $\alpha_i(T) = \frac{1}{2}h_i^\gamma(\mu_i^\gamma)^2$. It can be verified (see, e.g., [32]) that (45) satisfies the HJ-inequality for all $(t, x, u) \in [0, T] \times \mathbb{R} \times \mathbb{R}$.

Since the primary purpose of the example is to illustrate the characterization of the value function by (44), we restrict attention to the symmetric case where $\mu_1^d = -\mu_2^d = \mu_d$, $\sigma_1^d = \sigma_2^d$, and $x_0 = (\mu_1^\gamma + \mu_2^\gamma)/2 = 0$, thus $\rho_\gamma = 1/2$, and $\mu_1^\gamma = \mu^d$, and $\mu_2^\gamma = -\mu^d$. In particular, we consider the case with the desired distribution $\mathsf{p}_d = \frac{1}{2}\mathcal{N}(3, 1) + \frac{1}{2}\mathcal{N}(-3, 1)$ at $T = 2$, and hence we restrict attention to the sequence of functions parameterized by $\gamma = (\eta_\gamma, \frac{1}{2}, h_\gamma, 3, h_\gamma, -3) \equiv (\eta_\gamma, h_\gamma)$. The corresponding values of $v_\gamma(0, x_0) - \int_{\mathbb{R}^n} v_\gamma(T, x)$ are displayed over the region $(\eta_\gamma, h_\gamma) \in [1, 10^3] \times [0, 4]$ in Figure 1.

As observed in Figure 1a, for the family (45) of HJ-subsolutions, the supremum in not attained over the bounded domain and, while $h_{\gamma^*} = 0.97$, the suprimum requires $\eta_\gamma \to \infty$. Indeed, the value function in this case is nonsmooth and is required to be identified from $\lim_{(\eta_\gamma, h_\gamma)\to(\infty, 0.97)} v_\gamma(t, x)$ as displayed in Figure 1b. In order to illustrate that the desired probability distribution is attained, the optimal trajectories of 200 sample paths are displayed in Figure 1c and the empirical distribution of these trajectories obtained from 100‘000 sample paths are displayed in Figure 1d.

Acknowledgment

Discussions concerning this work with Panagiotis Tsiotras and Ramanarayan Vasudevan are gratefully acknowledged.

References

[1] A. Hotz and R. E. Skelton, “Covariance control theory,” *Int. J. Control*, vol. 46, no. 1, pp. 13–32, 1987.

[2] J.-H. Xu and R. Skelton, “An improved covariance assignment theory for discrete systems,” *IEEE Trans. Autom. Control*, vol. 37, no. 10, pp. 1588–1591, Oct. 1992.

[3] K. Yasuda, R. E. Skelton, and K. M. Grigoriadis, “Covariance controllers: A new parametrization of the class of all stabilizing controllers,” *Automatica*, vol. 29, no. 3, pp. 785–788, 1993.

[4] K. M. Grigoriadis and R. E. Skelton, “Minimum-energy covariance controllers,” *Automatica*, vol. 33, no. 4, pp. 569–578, 1997.

[5] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal steering of a linear stochastic system to a final probability distribution, part I,” *IEEE Trans. Autom. Control*, vol. 61, no. 5, pp. 1158–1169, May 2016.

[6] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal steering of a linear stochastic system to a final probability distribution, part II,” *IEEE Trans. Autom. Control*, vol. 61, no. 5, pp. 1170–1180, May 2016.

[7] A. Beghi, “On the relative entropy of discrete-time Markov processes with given end-point densities,” *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1529–1535, Sep. 1996.

[8] E. Bakolas, “Optimal covariance control for discrete-time stochastic linear systems subject to constraints,” in *Proc. 55th IEEE Conf. Decis. Control*, 2016, pp. 1153–1158.

[9] E. Bakolas, “Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints,” *Automatica*, vol. 91, pp. 61–68, May 2018.

[10] M. Goldshtein and P. Tsiotras, “Finite-horizon covariance control of linear time-varying systems,” in *Proc. 56th IEEE Conf. Decis. Control*, 2017, pp. 3606–3611.

[11] K. Okamoto, M. Goldshtein, and P. Tsiotras, “Optimal covariance control for stochastic systems under chance constraints,” *IEEE Control Syst. Lett.*, vol. 2, no. 2, pp. 266–271, Apr. 2018.

[12] K. Okamoto and P. Tsiotras, “Optimal stochastic vehicle path planning using covariance steering,” *IEEE Robot. Autom. Lett.*, vol. 4, no. 3, pp. 2276–2281, Jul. 2019.

[13] A. Halder and E. Wendel, “Finite horizon linear quadratic gaussian density regulator with Wasserstein terminal cost,” in *Proc. Amer. Control Conf.*, 2016, pp. 7249–7254.

[14] K. F. Caluya and A. Halder, “Finite horizon density control for static state feedback linearizable systems,” 2019, *arXiv:1904.02272*.

[15] J. Ridderhof, K. Okamoto, and P. Tsiotras, “Nonlinear uncertainty control with iterative covariance steering,” in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, 2019, pp. 3484–3490.

[16] A. Pakniyat and P. Tsiotras, “Steering the state of linear stochastic systems: A constrained minimum principle formulation,” in *Proc. IEEE Amer. Control Conf. (ACC)*, 2021, pp. 1300–1305.

[17] A. Pakniyat and P. Tsiotras, “Partially observed steering the state of linear stochastic systems,” in *Proc. 60th IEEE Conf. Decis. Control (CDC)*, 2021, pp. 3780–3785.

[18] R. B. Vinter and R. M. Lewis, “The equivalence of strong and weak formulations for certain problems in optimal control,” *SIAM J. Control Optim.*, vol. 16, no. 4, pp. 546–570, 1978.

[19] R. B. Vinter and R. M. Lewis, “A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumptions on the controls,” *SIAM J. Control Optim.*, vol. 16, no. 4, pp. 571–583, 1978.

[20] D. Vermes, “Optimal control of piecewise deterministic Markov processes,” *Stochast. Int. J. Probab. Stochast. Processes*, vol. 14, no. 3, pp. 165–207, 1985.

[21] W. H. Fleming and D. Vermes, “Convex duality approach to the optimal control of diffusions,” *SIAM J. control Optim.*, vol. 27, no. 5, pp. 1136–1155, 1989.

[22] H. Zhu, “Convex duality for finite-fuel problems in singular stochastic control,” *J. Optim. Theory Appl.*, vol. 75, no. 1, pp. 155–181, 1992.

[23] A. G. Bhatt and V. S. Borkar, “Occupation measures for controlled Markov processes: Characterization and optimality,” *Ann. Probab.*, vol. 24, no. 3, pp. 1531–1562, 1996.

[24] T. Kurtz and R. Stockbridge, “Existence of Markov controls and Characterization of optimal Markov controls,” *SIAM J. Control Optim.*, vol. 36, no. 2, pp. 609–653, 1998.

[25] M. Cho and R. Stockbridge, “Linear programming formulation for optimal stopping problems,” *SIAM J. Control Optim.*, vol. 40, no. 6, pp. 1965–1982, 2002.

[26] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, “Nonlinear optimal control via occupation measures and LMI-relaxations,” *SIAM J. control Optim.*, vol. 47, no. 4, pp. 1643–1666, 2008.

[27] E. Pauwels, D. Henrion, and J.-B. Lasserre, “Linear conic optimization for inverse optimal control,” *SIAM J. Control Optim.*, vol. 54, no. 3, pp. 1798–1825, 2016.

[28] Y. Chen and A. D. Ames, “Duality between density function and value function with applications in constrained optimal control and Markov decision process,” 2019, *arXiv:1902.09583*.

[29] R. Vasudevan, H. Gonzalez, R. Bajcsy, and S. S. Sastry, “Consistent approximations for the optimal control of constrained switched systems—Part 1: A conceptual algorithm,” *SIAM J. Control Optim.*, vol. 51, no. 6, pp. 4463–4483, 2013.

[30] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake, “Convex optimization of nonlinear feedback controllers via occupation measures,” *Int. J. Robot. Res.*, vol. 33, no. 9, pp. 1209–1230, 2014.

[31] P. Zhao, S. Mohan, and R. Vasudevan, “Optimal control of polynomial hybrid systems via convex relaxations,” *IEEE Trans. Autom. Control*, vol. 65, no. 5, pp. 2062–2077, May 2020.

[32] R. Salhab, R. P. Malhamé, and J. Le Ny, “Collective stochastic discrete choice problems: A min-LQG dynamic game formulation,” *IEEE Trans. Autom. Control*, vol. 65, no. 8, pp. 3302–3316, Aug. 2020.