

A Convex Approach to Stochastic Optimal Control Using Linear Operators

Umesh Vaidya and Bowen Huang

Abstract—The paper is about the optimal control of a stochastic dynamical system. We provide a convex formulation to the optimal control problem involving a stochastic dynamical system. The convex formulation is made possible by writing the stochastic optimal control problem in the dual space of densities involving the Fokker-Planck or Perron-Frobenius generator for a stochastic system. The convex formulation leads to an infinite-dimensional convex optimization problem for optimal control. We exploit Koopman and Perron-Frobenius generators' duality for the stochastic system to construct the finite-dimensional approximation of the infinite-dimensional convex problem. We present simulation results to demonstrate the application of the developed framework.

I. INTRODUCTION

The problem of optimal control for a nonlinear system in the presence of uncertainty is a classical problem in system theory with wide-ranging applications [1]. The existing approach for solving the stochastic optimal control problem (OCP) is based on dynamic programming, leading to a partial differential equation is known as Hamilton Jacobi Bellman (HJB) equation. The HJB equation is a nonlinear partial differential equation and is difficult to solve analytically, and one has to resort to the numerical solution of the HJB equation. This paper proposes a dual approach based on the density-based formulation of the optimal control problem. Duality is exploited in stochastic optimal control theory for the convex formulation of optimal transport problem [2], in the derivation of stochastic estimation problem [3], and Pontryagin principle for Ito equations [4]. The convex approach to the stochastic OCP problem presented in this paper is based on the duality in the stability theory of deterministic and stochastic dynamical systems [5]–[8]. This duality is exploited to provide convex formulation of stabilization problem for deterministic system

in [9]–[11] and deterministic optimal control problem [12]. Similarly, convex dual formulation involving occupation measures via moment-based approach for deterministic control problem is studied in [13]–[15]. In this paper, we extend the results from [12] to the stochastic setting. The duality in the stability theory at the fundamental level could be understood using duality between two linear operators in the dynamical system theory, namely Perron-Frobenius (P-F) and Koopman operator [16].

The main contributions of the paper are as follows. We provide a convex formulation to the infinite horizon stochastic OCP for a class of control affine nonlinear systems. Compared to the existing results on the convex formulation to the OCP in deterministic setting, we present the results using the theory of linear transfer operators. The convex formulation leads to infinite-dimensional convex optimization problem formulation of stochastic OCP. The results derived in this paper are based on duality between P-F and Koopman operators and their respective generators in a continuous-time stochastic setting. The P-F and Koopman generators in stochastic settings arise as Kolmogorov forward or Fokker-Planck equation and Kolmogorov backward equation. This connection of stochastic OCP with stochastic P-F and Koopman operators allows us to use methods discovered to approximate these operators for the finite-dimensional approximation of the infinite-dimensional optimization problem [17]–[20].

The paper is organized as follows. In Section II, we present preliminary results on the Koopman and P-F generators in the stochastic setting and almost everywhere stability results for stochastic systems. The main results on the convex formulation of stochastic OCP are presented in Section III. The finite-dimensional approximation of infinite-dimensional optimization problem is presented in Section IV followed by simulation results and conclusion in Sections V and VI respectively.

Financial support from National Science Foundation (NSF) Grant 2031573 is greatly acknowledged. U. Vaidya and B. Huang are with the Department of Mechanical Engineering, Clemson University, Clemson SC, USA, uvaidya@clemson.edu.

II. PRELIMINARIES AND NOTATIONS

In this section, we discuss some preliminaries and introduce some notations, which are used in deriving the main results on data-driven optimal control. Consider a dynamical system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sigma \mathbf{n}(\mathbf{x}) \xi, \quad \mathbf{x} \in \mathbb{R}^n. \quad (1)$$

$\mathbb{R}^q \ni \xi = \frac{dw}{dt}$ is the white noise process and is the time derivative of the Wiener process. We assume that $\mathbf{n}(0) = 0$, where the origin is assumed to be the system's equilibrium point. Let $\mathbf{X}_t$ be the solution of the stochastic differential equation (1). For more details on the definition and condition for the existence of a solution to a stochastic differential equation, refer to [16]. In particular, the vector fields $\mathbf{f}(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$ are assumed to satisfy following Lipschitz condition.

$$\begin{aligned} |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| &\leq L|\mathbf{x} - \mathbf{y}|, \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^n \\ |\mathbf{n}(\mathbf{x}) - \mathbf{n}(\mathbf{y})| &\leq L|\mathbf{x} - \mathbf{y}|, \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^n \end{aligned} \quad (2)$$

for some constant L . Let $\mathbf{X}^{\mathbf{x}} := \{\mathbf{X}_t^{\mathbf{x}}\}_{t \geq 0}$ be the solution process with initial condition $\mathbf{X}_0 = \mathbf{x}$, where the solution is defined in the sense of Ito calculus [16]. The solution is clearly the function of stochastic process, however for compactness of notation we do not denote this dependence. Let $P^{\mathbf{x}}$ be the distribution of $\mathbf{X}^{\mathbf{x}}$ and $\mathbb{E}_{\mathbf{x}}$ be the expectation with respect to $P^{\mathbf{x}}$. We introduce following notations. Let $\mathcal{L}_1(\mathbb{R}^n)$, integrable functions on $\mathbb{R}^n$, $\mathcal{L}_{\infty}(\mathbb{R}^n)$, bounded functions on $\mathbb{R}^n$, $\mathcal{C}^k(\mathbb{R}^n)$ space of functions with k continuous derivative, $\mathcal{C}_c^k(\mathbb{R}^n)$, space of functions in $\mathcal{C}^k(\mathbb{R}^n)$ with compact support, and $\mathcal{C}_0$ closure of $\mathcal{C}_c^k(\mathbb{R}^n)$ in $\mathcal{L}_{\infty}$ norm.

A. Perron-Frobenius and Koopman Operator for Stochastic System

The results and discussion in this section is taken from various references [16], [21]–[23]. The linear operators' theory involving Perron-Frobenius (P-F) and Koopman operator from deterministic dynamical system generalizes to the stochastic system.

Following assumptions are made on the vector field $\mathbf{f}$ and $\mathbf{n}$ to ensure that these linear operators and their solutions are well defined.

Assumption 1: We assume that the vector field, $\mathbf{f}(\mathbf{x})$, and coefficients $a_{ij}(\mathbf{x}) = \sum_{k=1}^q \mathbf{n}_{ik}(\mathbf{x}) \mathbf{n}_{kj}(\mathbf{x})$ are $\mathcal{C}^4$ functions of $\mathbf{x}$. Furthermore, following growth conditions are satisfied by the coefficients

$$|a_{ij}(\mathbf{x})| \leq M(1 + |\mathbf{x}|^2), |\tilde{f}_i(\mathbf{x})| \leq M(1 + |\mathbf{x}|), |\tilde{c}(\mathbf{x})| \leq M,$$

$i, j = 1, \dots, n$, where M is some positive constant, $\mathbf{f} = (f_1, \dots, f_n)^{\top}$ and

$$\tilde{f}_i(\mathbf{x}) = -f_i(\mathbf{x}) + \sigma^2 \sum_{j=1}^n \frac{\partial a_{ij}(\mathbf{x})}{\partial x_j}, \quad (3)$$

$$\tilde{c}(\mathbf{x}) = \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2 a_{ij}(\mathbf{x})}{\partial x_i \partial x_j} - \sum_{i=1}^n \frac{\partial f_i(\mathbf{x})}{\partial x_i} \quad (4)$$

We have $a_{ij} = a_{ji}$. For any given $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$, we assume that uniform parabolicity assumptions holds if there exists a constant $\rho > 0$ such that

$$\sum_{i,j=1}^n a_{ij} \lambda_i \lambda_j \geq \rho \sum_{i=1}^n \lambda_i^2$$

Definition 1 (Koopman Operator and its Generator): Let $\varphi(\mathbf{x}) \in \mathcal{L}_{\infty}(\mathbb{R}^n)$, then the Koopman operator $\mathbb{U}_t : \mathcal{L}_{\infty}(\mathbb{R}^n) \rightarrow \mathcal{L}_{\infty}(\mathbb{R}^n)$ for stochastic dynamical system (1) is defined as

$$[\mathbb{U}_t \varphi](\mathbf{x}) = \mathbb{E}_{\mathbf{x}} [\varphi(\mathbf{X}_t^{\mathbf{x}})] \quad (5)$$

$\mathbb{U} := \{\mathbb{U}_t\}_{t \geq 0}$ restricts to strongly continuous semi-group on $\mathcal{C}_0$ and hence it has a infinitesimal generator ([24], Theorem 21.11). For any $\varphi \in \mathcal{C}_c^2(\mathbb{R}^n)$, the infinitesimal generator is given by

$$\mathcal{A}_K \varphi := \lim_{t \rightarrow 0} \frac{(\mathbb{U}_t - I) \varphi}{t} \quad (6)$$

$$= \mathbf{f}(\mathbf{x}) \cdot \nabla \varphi + \frac{\sigma^2}{2} \sum_{i,j=1}^n [\mathbf{n} \mathbf{n}^{\top}]_{ij} \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \quad (7)$$

We assume that the distribution of $\mathbf{X}_0$ is absolutely continuous and has density $\rho_0(\mathbf{x})$. Then we know that $\mathbf{X}_t$ has a density $\rho(\mathbf{x}, t)$ i.e.,

$$\text{Prob}\{\mathbf{X}_t \in B\} = \int_B \rho(\mathbf{x}, t) d\mathbf{x}$$

which satisfies the following Fokker-Planck (F-P) equation also known as Kolomogorov forward equation

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} = -\nabla \cdot (\mathbf{f}(\mathbf{x}) \rho) + \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2 [(\mathbf{n} \mathbf{n}^{\top})_{ij} \rho]}{\partial x_i \partial x_j} \quad (8)$$

Following Assumption 1, we know the solution $\rho(\mathbf{x}, t)$ to F-P equation exists and is differentiable (Theorem 11.6.1 [16]). Under some regularity assumptions on the coefficients of the F-P equation (Definition 11.7.6 [16]) it can be shown that the F-P admits a unique classical solution (Theorem 11.7.1 [16]). The classical solution is used to define family of operators. Given any initial density function $\rho(\mathbf{x}, 0) = h(\mathbf{x}) \in \mathcal{L}_1(\mathbb{R}^n)$, we can define family of operators $\{\mathbb{P}_t\}_{t \geq 0}$ by

$$\rho(\mathbf{x}, t) = [\mathbb{P}_t h](\mathbf{x}) = \int \Gamma(\mathbf{x}, \mathbf{y}, t) h(\mathbf{y}) d\mathbf{y} \quad (9)$$

Definition 2 (P-F Operator and Generator): The family of operators $\mathbb{P} := \{\mathbb{P}_t\}_{t \geq 0}$ as defined in (9) are called stochastic Perron-Frobenius semi-group. Following [16] (Theorems 11.6.1, 11.7.1, Corollary 11.8.1, and Remark 11.8.1) it also follows that the right hand side of the F-P equation is also the infinitesimal generator for the stochastic P-F semi-group. In particular, we have

$$\begin{aligned} \mathcal{A}_{PF}\psi &:= \lim_{t \rightarrow 0} \frac{(\mathbb{P}_t - I)\psi}{t} \\ &= -\nabla \cdot (\mathbf{f}(\mathbf{x})\psi) + \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2 [(\mathbf{n}\mathbf{n}^\top)_{ij}\psi]}{\partial x_i \partial x_j} \end{aligned} \quad (10)$$

Since $\mathbb{P}_t$ is a semi-group with generator $\mathcal{A}_{PF}$,

$$\frac{d}{dt} \mathbb{P}_t \psi = \mathcal{A}_{PF} \mathbb{P}_t \psi \quad (11)$$

The duality between the P-F and Koopman generators can be expressed as follows ([16] Theorem 11.6.1):

$$\int_{\mathbb{R}^n} [\mathcal{A}_K \varphi](\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \varphi(\mathbf{x}) [\mathcal{A}_{PF} \psi](\mathbf{x}) d\mathbf{x} \quad (12)$$

The duality relation between the two semi-groups can be expressed as follows (We refer the readers to [25] for details of this proof)

$$\int_{\mathbb{R}^n} [\mathbb{U}_t \varphi](\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \varphi(\mathbf{x}) [\mathbb{P}_t \psi](\mathbf{x}) d\mathbf{x} \quad (13)$$

Definition 3: [Stochastic Almost everywhere stability] The equilibrium point at $x = 0$ is said to be almost everywhere almost sure stable with respect to finite measure $\mu \in \mathcal{M}(\mathbb{R}^n)$ if for almost all initial condition $\mathbf{x}$ w.r.t. measure μ , we have

$$\text{Prob}\{\lim_{t \rightarrow \infty} \mathbf{X}_t^{\mathbf{x}} = 0\} = 1 \quad (14)$$

III. CONVEX FORMULATION OF STOCHASTIC OCP

In this section we present the main results of this paper on the convex formulation to the optimal control problem. We consider optimal control problem for affine in control dynamical system of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \sigma \mathbf{n}(\mathbf{x})\xi \quad (15)$$

where, $\mathbf{x} \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}$ is the control input. For the simplicity of presentation we present the results for the case of single input, however the results generalize to multi-input case in a straight forward manner. We make following assumption.

Assumption 2: We assume that the $\mathbf{x} = 0$ is locally almost sure asymptotically stable equilibrium point for system (15) with control input $u = 0$. In particular,

there exists a neighborhood $\mathcal{N}$ of the origin such that for all $\mathbf{x} \in \mathcal{N}$,

$$\text{Prob}\{\mathbf{X}_t^{\mathbf{x}} \in \mathcal{N}\} = 1 \quad \forall t \geq 0, \quad \& \quad (16)$$

$$\text{Prob}\{\lim_{t \rightarrow \infty} \mathbf{X}_t^{\mathbf{x}} = 0\} = 1. \quad (17)$$

Remark 1: From the computational perspective local stability assumption is not needed as local stabilizing controller can always be designed by assuming local stabilizability assumption on the linear dynamics meaning the pair $(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(0), \mathbf{g}(0))$.

We denote by $\mathbf{S}_1 := \mathbb{R}^n \setminus \mathcal{N}$, where $\mathcal{N}$ is the neighborhood of the origin as defined in Assumption 2.

Consider the cost function

$$J(\mu) = \int_{\mathbf{S}_1} \mathbb{E}_{\mathbf{x}} \left[\int_0^\infty q(\mathbf{X}_t^{\mathbf{x}}) + ru(t)^2 dt \right] d\mu(\mathbf{x}) \quad (18)$$

Unlike cost function in the usual formulation of stochastic OCP i.e., $V(\mathbf{x}) = \mathbb{E}_{\mathbf{x}} \left[\int_0^\infty q(\mathbf{X}_t^{\mathbf{x}}) + ru^2 dt \right]$, where the expectation are taken only over the different realization of random processes, in our proposed OCP the expectation are also taken w.r.t. measure μ distribution over initial conditions i.e., $J = \int_{\mathbf{S}_1} V(\mathbf{x}) d\mu(\mathbf{x})$.

Following assumption are made on the state cost function $q(\mathbf{x})$ and the measure μ .

Assumption 3: We assume that the state cost function $q : \mathbb{R}^n \rightarrow \mathbb{R}^+$ is zero at the origin and uniformly bounded away from zero outside the neighborhood $\mathcal{N}$. The measure $\mu \in \mathcal{M}(\mathbb{R}^n)$ is equivalent to Lebesgue with Radon-Nikodým derivative h i.e., $\frac{d\mu}{dx} = h(\mathbf{x}) > 0$ and $h \in \mathcal{L}_1(\mathbb{R}^n, \mathbb{R}_{>0}) \cap \mathcal{C}^2(\mathbb{R}^n)$.

Following important assumption is made on the solution of the optimal control problem.

Assumption 4: We assume that there exists a feedback control such that the cost function with this control is finite and the controller is assumed to be stochastic a.e stabilizing the control system. Furthermore, the optimal control is feedback in nature i.e., $u = k(\mathbf{x})$ with the function k is assumed to be $\mathcal{C}^4$ function of $\mathbf{x}$.

Remark 2: The above assumption on the stochastic a.e. stabilizing property of the feedback control input is restrictive and is made to simplify the proof of the main result and due to space constraint. In [25], we prove the main result without this restrictive assumption.

With the assumed feedback form of the control input, the OCP can be written as

$$\begin{aligned} \inf_{k \in \mathcal{C}^4} \int_{\mathbf{S}_1} \mathbb{E}_{\mathbf{x}} \left[\int_0^\infty q(\mathbf{X}_t^{\mathbf{x}}) + rk(\mathbf{X}_t^{\mathbf{x}})^2 dt \right] d\mu(\mathbf{x}) \\ \text{s.t.} \quad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})k(\mathbf{x}) + \mathbf{n}(\mathbf{x})\xi \end{aligned} \quad (19)$$

Following is one of the main results of this paper.

Theorem 3: Consider the optimal control problem (19), with the system dynamics, cost function, and optimal control satisfying Assumptions 2, 3, and 4 respectively. The OCP (19) can be written as following infinite dimensional convex optimization problem

$$\begin{aligned} J^*(\mu) &= \inf_{\rho \in \mathcal{S}_1, \bar{\rho} \in \mathcal{C}^4} \int_{\mathbf{S}_1} q(\mathbf{x})\rho(\mathbf{x}) + r \frac{\bar{\rho}(\mathbf{x})^2}{\rho} d\mathbf{x} \\ \text{s.t. } \nabla \cdot (\mathbf{f}\rho + \mathbf{g}\bar{\rho}) - \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2[(\mathbf{nn}^\top)_{ij}\rho]}{\partial \mathbf{x}_i \partial \mathbf{x}_j} &= h. \end{aligned} \quad (20)$$

where $\mathcal{S} := \mathcal{L}_1(\mathbf{S}_1) \cap \mathcal{C}^2(\mathbf{S}_1, \mathbb{R}_{\geq 0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as $k(\mathbf{x}) = \frac{\bar{\rho}(\mathbf{x})}{\rho(\mathbf{x})}$. Furthermore, the optimal control $k(\mathbf{x})$ is almost everywhere uniform stochastic stabilizing the origin.

Proof: We can move the expectation inside the integral of (19) to write $J(\mu)$ as

$$J(\mu) = \int_{\mathbf{S}_1} \int_0^\infty \mathbb{E}_{\mathbf{x}} \left[q(\mathbf{X}_t^{\mathbf{x}}) + \mathbf{u}_t^\top \mathbf{R} \mathbf{u}_t \right] dt h(\mathbf{x}) d\mathbf{x}$$

Using Assumption (4) on the feedback property of the optimal input and using $\varphi(\mathbf{x}) := q(\mathbf{x}) + rk(\mathbf{x})^2$, we obtain using the definition of Koopman operator

$$J(\mu) = \int_{\mathbf{S}_1} \int_0^\infty [\mathbb{U}_t^c \varphi](\mathbf{x}) dt h(\mathbf{x}) d\mathbf{x}$$

where $\mathbb{U}_t^c$ and $\mathbb{P}_t^c$ are notation for the Koopman and P-F semi-groups for the closed loop system $\mathbf{f}_c := \mathbf{f} + \mathbf{g}k + \sigma \mathbf{n}\xi$. Using the linear property of the Koopman operator and the duality of Koopman and P-F semi-groups we obtain

$$J = \int_{\mathbf{S}_1} \int_0^\infty \varphi(\mathbf{x}) [\mathbb{P}_t^c h](\mathbf{x}) dt d\mathbf{x} \quad (21)$$

where we define $\rho(\mathbf{x}) := \int_0^\infty [\mathbb{P}_t^c h](\mathbf{x}) dt$. We next show that $\rho(\mathbf{x})$ is well defined for almost all $\mathbf{x}$. Using the Assumption 3 on the uniform lower bound, say κ , for state cost function outside $\mathcal{N}$, and Assumption 4 on the finite value of cost, we obtain

$$\kappa \int_{\mathbf{S}_1} \rho(\mathbf{x}) d\mathbf{x} \leq \int_{\mathbf{S}_1} \varphi(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x} \leq M < \infty$$

This show that $\rho(\mathbf{x})$ is integrable function on $\mathbf{S}_1$. We next show that $\rho(\mathbf{x})$ satisfy the following equation

$$\nabla \cdot ((\mathbf{f}_c(\mathbf{x})\rho(\mathbf{x})) - \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2[(\mathbf{nn}^\top)_{ij}\rho]}{\partial \mathbf{x}_i \partial \mathbf{x}_j}) = h(\mathbf{x}), \quad (22)$$

for a.e. $\mathbf{x} \in \mathbf{S}_1$. Substituting the integral formula for $\rho(\mathbf{x})$ in (22), we obtain

$$\begin{aligned} &\nabla \cdot (\mathbf{f}_c(\mathbf{x})\rho(\mathbf{x})) - \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2[(\mathbf{nn}^\top)_{ij}\rho]}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \\ &= \int_0^\infty \nabla \cdot (\mathbf{f}_c(\mathbf{x})[\mathbb{P}_t^c h](\mathbf{x})) dt \\ &- \int_0^\infty \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2[(\mathbf{nn}^\top)_{ij}[\mathbb{P}_t^c h](\mathbf{x})]}{\partial \mathbf{x}_i \partial \mathbf{x}_j} dt \\ &= \int_0^\infty -\frac{d}{dt} [\mathbb{P}_t^c h](\mathbf{x}) dt = -[\mathbb{P}_t^c h](\mathbf{x}) \Big|_{t=0}^\infty = h(\mathbf{x}) \end{aligned} \quad (23)$$

where we have used the infinitesimal generator property of P-F operator Eq. (11) and the fact that $\lim_{t \rightarrow \infty} [\mathbb{P}_t^c h](\mathbf{x}) = 0$ (Assumption 4). Furthermore, since $h > 0$, it follows that $\rho > 0$ from the positivity property of P-F semigroup $\mathbb{P}_t^c$. Combining (21) and (24) and by defining $\bar{\rho}(\mathbf{x}) := \rho(\mathbf{x})k(\mathbf{x})$, it follows that the OCP problem can be written as convex optimization problem (20). The optimal solution $\rho^*(\mathbf{x}) \in \mathcal{L}_1(\mathbf{S}_1) \cap \mathcal{C}^2(\mathbf{S}_1, \mathbb{R}_{\geq 0})$ follows from the fact that $h \in \mathcal{L}_1(\mathbb{R}^n, \mathbb{R}_{>0}) \cap \mathcal{C}^2(\mathbb{R}^n)$ from Assumption 3 and the definitions of ρ (integral formula) and P-F operator (9). $\blacksquare$

IV. COMPUTATION OF STOCHASTIC OPTIMAL CONTROL

We first discuss results on the approximation of the Koopman and P-F generators.

A. Approximation of Koopman and P-F Generators for Stochastic System

For the approximation of the Koopman and P-F generators we use Naturally Structured Dynamic Mode Decomposition (NSDMD) algorithm from [17]. We first outline the procedure for constructing the approximation of P-F generator. Let $\{\mathbf{x}_i, \mathbf{y}_i^\ell\}$ be the two consecutive snapshots obtained by simulating the system (1) using Euler-Maruyama discretization. In particular, we have

$$\mathbf{y}_i^\ell = \mathbf{x}_i + \Delta t \mathbf{f}(\mathbf{x}_i) + \sigma \sqrt{\Delta t} \mathbf{n}(\mathbf{x}_i) \xi^\ell \quad (24)$$

where $i = 1, \dots, N$ are the number of initial conditions assume to be uniformly distributed in the state space and $\ell = 1, \dots, R$ are the number of realization. For the finite dimensional approximation, let $\Psi : \mathbf{X} \rightarrow \mathbb{R}^K$ be the vector valued positive basis functions. Writing

$$\phi(\mathbf{x}) = \Psi(\mathbf{x})^\top \mathbf{a}, \quad \hat{\phi}(\mathbf{x}) = \Psi(\mathbf{x})^\top \hat{\mathbf{a}}$$

For the approximation of the Koopman operator in Eq. (5) the expectation is replaced with empirical mean.

$$\hat{\phi}(\mathbf{x}_i) = \Psi^\top(\mathbf{x}_i) \hat{\mathbf{a}} = [\mathbb{U}_t \phi](\mathbf{x}_i) + r \quad (25)$$

$$= \mathbb{E}_{\mathbf{x}}[\phi(\mathbf{X}_{\Delta t}^{\mathbf{x}})] + r \approx \frac{1}{R} \sum_{\ell=1}^R \Psi^\top(\mathbf{y}_i^\ell)^\top \mathbf{a} + r \quad (26)$$

The objective is to minimize the residual term r . The approximation of Koopman operator is written as

$$\begin{aligned} \min_{\mathbf{K} \in \mathbb{R}^{K \times K}} \quad & \|\mathbf{G}\mathbf{K} - \mathbf{A}\|_F \\ \text{s.t.} \quad & [\mathbf{\Lambda}\mathbf{K}\mathbf{\Lambda}^{-1}]_{ij} \geq 0, \quad \mathbf{\Lambda}\mathbf{K}\mathbf{\Lambda}^{-1}\mathbf{1} = \mathbf{1} \end{aligned} \quad (27)$$

where

$$\mathbf{G} = \frac{1}{M} \sum_{i=1}^N \Psi(\mathbf{x}_i) \Psi^\top(\mathbf{x}_i), \mathbf{A} = \frac{1}{MR} \sum_{i,\ell}^{N,R} \Psi(\mathbf{x}_i) \Psi^\top(\mathbf{y}_i^\ell)$$

and $\mathbf{\Lambda} = \int \Psi(\mathbf{x}) \Psi^\top(\mathbf{x}) d\mathbf{x}$. The constraints are imposed to ensure that the resulting matrix $\mathbf{K}$ is positive and satisfy the Markov property. The P-F operator and the generator for the stochastic system is then obtained using duality relation [17] as

$$\mathbb{P}_{\Delta t} \approx \mathbf{P} = \mathbf{\Lambda}^{-1} \mathbf{K}^\top \mathbf{\Lambda}, \quad \mathcal{A}_{PF}^{f+\sigma n} \approx \frac{\mathbf{P} - \mathbf{I}}{\Delta t} =: \mathbf{M}_0$$

For the approximation of P-F generator corresponding to the control vector field i.e., $\mathbf{g}(\mathbf{x})$, we generate the data from the deterministic dynamical system $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ i.e., $\mathbf{y}_i = \mathbf{x}_i + \Delta t \mathbf{g}(\mathbf{x}_i)$. Now using the above procedure outlined for the approximation of the P-F generator for stochastic system we approximate the P-F generator for vector field $\mathbf{g}$. The only difference is, since the system $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ is deterministic we do not generate multiple realization of trajectories. We denote by $\mathbf{M}_1$, the P-F generator corresponding to the vector field $\mathbf{g}$.

B. Approximation of the Optimization problem

We begin with the approximation of the constraints in the optimization problem.

Assumption 5: We assume that the basis functions, $\psi_k(\mathbf{x})$ for $k = 1, \dots, N$ are positive and let

$$\Psi(\mathbf{x}) = [\psi_1(\mathbf{x}), \dots, \psi_N(\mathbf{x})]^\top.$$

In this paper, we use Gaussian RBF to obtain all the simulation results i.e., $\psi_k(\mathbf{x}) = \exp^{-\frac{\|\mathbf{x} - \mathbf{c}_k\|^2}{\sigma^2}}$. where $\mathbf{c}_k$ is the center of the k^{th} Gaussian RBF.

Let $\rho(\mathbf{x})$, and $\bar{\rho}(\mathbf{x})$ be expressed as follows

$$\rho(\mathbf{x}) \approx \Psi^\top \mathbf{v}, \quad \bar{\rho}(\mathbf{x}) \approx \Psi^\top \mathbf{w}, \quad h(\mathbf{x}) = \Psi^\top \mathbf{m} \quad (28)$$

Using the procedure outlined in Section (IV-A) for the approximation of the P-F generators for the stochastic system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sigma \mathbf{n}(\mathbf{x}) \xi$ and $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$, we obtain following approximation for the constraints

$$\nabla \cdot (\mathbf{f}\rho) - \frac{\sigma^2}{2} \sum_{i,j=1}^n \frac{\partial^2 [(\mathbf{n}\mathbf{n}^\top)_{ij} \rho]}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \approx -\Psi(\mathbf{x})^\top \mathbf{M}_0 \mathbf{v}$$

$$\nabla \cdot (\mathbf{g}\bar{\rho}) \approx -\Psi(\mathbf{x})^\top \mathbf{M}_1 \mathbf{w}$$

We now proceed with the approximation of the cost

$$\int_{\mathbf{x}_1} q(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x} \approx \int_{\mathbf{x}} q(\mathbf{x}) \Psi^\top d\mathbf{x} \mathbf{v} = \mathbf{d}^\top \mathbf{v}$$

where the vector $\mathbf{d} := \int_{\mathbf{x}} q(\mathbf{x}) \Psi^\top d\mathbf{x}$ is pre-computed. Similarly, we approximate $\frac{\bar{\rho}}{\rho} = \Psi^\top \frac{\mathbf{w}}{\mathbf{v}}$, where we assume element-wise division, hence

$$\frac{\bar{\rho}^2}{\rho} \approx \mathbf{w}^\top \Psi \Psi^\top \frac{\mathbf{w}}{\mathbf{v}}, \quad \int_{\mathbf{x}_1} r \frac{\bar{\rho}^2}{\rho} d\mathbf{x} \approx \mathbf{w}^\top \mathbf{D} \frac{\mathbf{w}}{\mathbf{v}}$$

where, $\mathbf{D} = \int_{\mathbf{x}_1} \Psi \Psi^\top d\mathbf{x}$. We have the following approximation to the optimization problem (20)

$$\begin{aligned} \min_{\Psi^\top \mathbf{v} \geq 0, \mathbf{w}} \quad & \mathbf{d}^\top \mathbf{v} + r \mathbf{w}^\top \mathbf{D} \frac{\mathbf{w}}{\mathbf{v}} \\ \text{s.t.} \quad & -\Psi(\mathbf{x})^\top (\mathbf{M}_0 \mathbf{v} + \mathbf{M}_1 \mathbf{w}) = \Psi(\mathbf{x})^\top \mathbf{m} \end{aligned}$$

Since the basis functions are assumed to be positive (Assumption 5), the approximation of ρ and $\bar{\rho}$ in (28) is obtained by solving the following problem.

$$\begin{aligned} \min_{\mathbf{v} \geq 0, \mathbf{w}} \quad & \mathbf{d}^\top \mathbf{v} + r \mathbf{w}^\top \mathbf{D} \frac{\mathbf{w}}{\mathbf{v}} \\ \text{s.t.} \quad & -(\mathbf{M}_0 \mathbf{v} + \mathbf{M}_1 \mathbf{w}) = \mathbf{m} \end{aligned} \quad (29)$$

The optimal control is then approximated as $u = \Psi^\top(\mathbf{x}) \frac{\mathbf{w}}{\mathbf{v}}$, where the division is element-wise.

V. SIMULATION RESULTS

Controlled Van der Pol oscillator

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = (1 - x_1^2)x_2 - x_1 + \sigma x_1 \xi_2 + u \quad (30)$$

where $x \in \mathbb{R}^2$ and $u \in \mathbb{R}$ is the single input. For this example we assume $q(\mathbf{x}) = \mathbf{x}^\top \mathbf{x}$ and quadratic cost on control. The finite dimensional optimization formulation in Eq. (29) is applied for the design of optimal control. For the approximation of stochastic P-F operator, we applied NSDMD algorithm using one-step time-series data with 10000 initial conditions, $\Delta t = 0.01$ (i.e., 10^4 data samples). We used 225 Gaussian RBF with $\bar{\sigma} = 0.2$, and the centers of RBF are distributed uniformly in $D = [-2, 2] \times [-3, 3]$. The noise $\sigma = 0.5$. In Fig. 1 and Fig. 2 we show the successful simulation results for the comparison of the open/ closed loop trajectories from five initial conditions. The simulation results are performed using MATLAB on desktop machine with 16 GB RAM. The total simulation time is less than 5 min.

VI. CONCLUSION

This paper presented results on the convex formulation of the stochastic optimal control problem leading to an infinite-dimensional convex optimization problem. The data-driven methods developed for the finite-dimensional approximation of Koopman and P-F

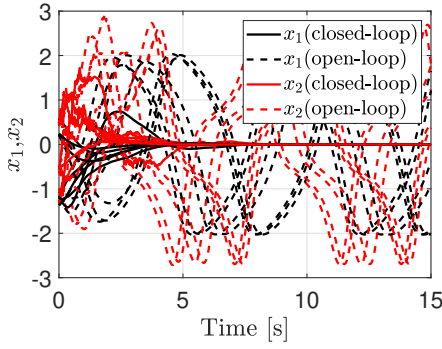


Fig. 1: $x_{1\sim 2}$ vs t for Van der pol

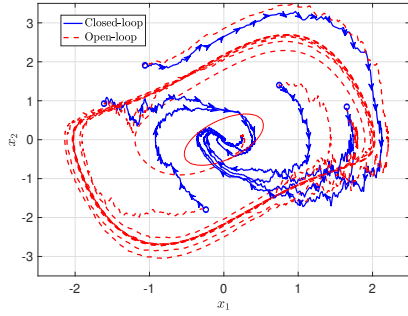


Fig. 2: Trajectories in 2-D space

operators are used in the finite-dimensional approximation of the optimization problem. Future research efforts will focus on discovering data-driven methods that do not involve knowing the system model for stochastic optimal control computation. Also, the choice of positive Gaussian RBF in the approximation process is restrictive. Future research efforts will focus on relaxing this assumption.

REFERENCES

- [1] R. F. Stengel, "Stochastic optimal control: theory and application." *New York*, 1986.
- [2] T. Mikami and M. Thieullen, "Duality theorem for the stochastic optimal control problem," *Stochastic processes and their applications*, vol. 116, no. 12, pp. 1815–1835, 2006.
- [3] E. Todorov, "General duality between optimal control and estimation," in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 4286–4292.
- [4] J.-M. Bismut, "An introductory approach to duality in optimal stochastic control," *SIAM review*, vol. 20, no. 1, pp. 62–78, 1978.
- [5] A. Rantzer, "A dual to Lyapunov's stability theorem," *Systems & Control Letters*, vol. 42, pp. 161–168, 2001.
- [6] U. Vaidya and P. G. Mehta, "Lyapunov measure for almost everywhere stability," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 307–323, 2008.
- [7] R. Van Handel, "Almost global stochastic stability," *SIAM journal on control and optimization*, vol. 45, no. 4, pp. 1297–1313, 2006.
- [8] U. Vaidya, "Stochastic stability analysis of discrete-time system using lyapunov measure," in *2015 American Control Conference (ACC)*. IEEE, 2015, pp. 4646–4651.
- [9] S. Prajna, P. A. Parrilo, and A. Rantzer, "Nonlinear control synthesis by convex optimization," *IEEE Transactions on Automatic Control*, vol. 49, no. 2, pp. 1–5, 2004.
- [10] U. Vaidya, P. Mehta, and U. Shanbhag, "Nonlinear stabilization via control lyapunov measure," *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1314–1328, 2010.
- [11] H. Choi, U. Vaidya, and Y. Chen, "A convex data-driven approach for nonlinear control synthesis," *arXiv preprint arXiv:2006.15477*, 2020.
- [12] B. Huang and U. Vaidya, "A convex approach to data-driven optimal control via perron-frobenius and koopman operators," *Under review*, 2020.
- [13] D. Henrion and M. Korda, "Convex computation of the region of attraction of polynomial control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 297–312, 2013.
- [14] M. Korda, "Moment-sum-of-squares hierarchies for set approximation and optimal control," IGM, Lausanne, Tech. Rep., 2016.
- [15] M. Korda, D. Henrion, and C. N. Jones, "Convergence rates of moment-sum-of-squares hierarchies for optimal control problems," *Systems & Control Letters*, vol. 100, pp. 1–5, 2017.
- [16] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. New York: Springer-Verlag, 1994.
- [17] B. Huang and U. Vaidya, "Data-driven approximation of transfer operators: Naturally structured dynamic mode decomposition," in *2018 American Control Conference*. IEEE, 2018, pp. 5659–5664.
- [18] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.
- [19] M. Korda and I. Mezić, "On convergence of extended dynamic mode decomposition to the koopman operator," *Journal of Nonlinear Science*, vol. 28, no. 2, pp. 687–710, 2018.
- [20] S. Klus, F. Nüske, S. Peitz, J.-H. Niemann, C. Clementi, and C. Schütte, "Data-driven approximation of the koopman generator: Model reduction, system identification, and control," *Physica D: Nonlinear Phenomena*, vol. 406, p. 132416, 2020.
- [21] Y. Shi and B. Liu, "Fokker-planck equation for kolmogorov operators associated to stochastic pde with multiplicative noise," *Advances in Difference Equations*, vol. 2014, no. 1, p. 222, 2014.
- [22] L. Manca, "Kolmogorov equations for measures," *Journal of Evolution Equations*, vol. 8, no. 2, pp. 231–262, 2008.
- [23] R. Rudnicki, K. Pichór, and M. Tyran-Kamińska, "Markov semigroups and their applications," in *Dynamics of Dissipation*. Springer, 2002, pp. 215–238.
- [24] O. Kallenberg, *Foundations of modern probability*. Springer Science & Business Media, 2006.
- [25] U. Vaidya and D. Castro, "Data-driven convex approach to stochastic optimal control using linear operators," Submitted for Publication, Tech. Rep., 2021.

Comments

This article proposes controller synthesis methods using the Perron-Frobenius and Koopman operators for control affine, nonlinear, stochastic systems. The reviewers agree that the contributions are important, but would benefit from further discussion of assumptions made. Contextualization compared to the authors' previous work and comparison to standard methods would highlight the contributions. Several notational, grammatical, and typographical errors were noted by each reviewer.

Reviewer 1

The following are the comments in order: - Eq. (2) - Is it the same Lipschitz constant for both f and n ? - The way X^x is defined after Eq. (2), it appears deterministic as it is not a function of the stochastic process - The set B after Eq. (7) is not defined - α in the paragraph before Eq. (9) is not defined - After Eq. (30), the function g seems to be misused. In Eq. (1), n is used whereas here, g is used in the Eq. $\dot{x} = f(x) + \sigma g(x)\xi$ - Is it assumed that the control input is single-dimensional? - The legend in Figures 1 and 3 is not consistent and the trajectories for many initial conditions is shown which is difficult to see - Figures 2 and 4, it is not clear where is the initial condition as the direction of the trajectories is not shown It is recommended the authors correct the typos and grammatical errors in the manuscript (for example, in references, 'koopman' $\rightarrow$ 'Koopman')

Reviewer 2

The paper considers optimal control synthesis of stochastic dynamical systems and proposes a convex optimization based approach for control-affine systems. This is made possible by exploiting duality and using finite dimensional approximation of the associated Koopman/PF operators. The manuscript is well organized and presents novel theoretical contributions. I have a few comments for the authors that need elaboration/justification

- The theoretical contributions presented in the paper hinge on some strong and restrictive assumptions. I believe the authors should provide further elaboration to help address their practical implications

- For instance, it is assumed that the functions ρ and $\bar{\rho}$ lie in the span of the basis functions such that they can be expressed as a linear combination which further results in the optimization problem to solve for coefficients v, w instead of the functions. This is a very contested topic in the operator theoretic research community. Could the authors comment on its validity?

- It is also assumed that the basis functions are positive. This rather restricts the choice of basis functions for practical implementation. RBFs may work for simple systems like Van Der Pol oscillators but fail for other systems.

- It is also curious to see how the method performs for different levels of stochasticity. The authors can add more results with different levels of random noise terms.

Reviewer 3

his paper proposes a convex approach of the stochastic optimal control problem for a class of control affine nonlinear systems. The stochastic optimal control problem is formulated as an infinite-dimensional convex optimization problem by the density-based formulation. Then, the duality between Koopman and Perron-Frobenius generators for the stochastic system is exploited to construct the finite-dimensional approximation of the infinite-dimensional linear program. This paper is well written and the numerical examples properly shows the effectiveness of the proposed method.

1. The cost in (21) is weighted by measure Image for each Image. It would be better to provide detailed explanation about the meaning and ground of this form of cost function.
2. The detailed procedure of the

approximation for cost function and constraints is not provided in section IV. B. It would be better to provide more detailed derivation process for them. 3. Please check (9) and the equation above (9). They don't match. 4. Below (20): The definition of Image is not provided. 5. Section III.A : Image seems to omit the term Image. 6. Section III.A : Definition for Image is not provided. 7. (24): Definition for Image is not provided. 8. (28): Image is used repeatedly on the right-hand side of (28). 9. Above (27): Definitions for Image, Image are not provided. 10. Equation below (29): Please check if M in the denominator should be N.

Reviewer 4

In this article, recent results by the authors on feedback control synthesis using the Perron- Frobenius and Koopman operator (A Convex Approach to Data-driven Optimal Control via Perron- Frobenius and Koopman Operators, arXiv:2010.01742) are extended to stochastic systems. Under several assumptions, the numerical approximation of the Koopman operator via Extended Dynamic Mode Decomposition is used to compute a density function based on which a feedback signal can be computed that yields a dual Lyapunov function and thus stabilizes the system. Two example systems are used to validate the results. The paper covers an interesting topic and the results are certainly of great interest for the control community. Even though there is a significant overlap with the paper on which this extension is based, the novelty regarding stochastic systems is sufficient for a new paper. I thus suggest to accept it for publication provided that a few points are addressed. • Can you comment on Assumption 1? Is it restrictive or very common for stochastic dynamical systems? • The matrix n in (1) is not introduced properly. • How does the performance of the feedback controller compare to established methods? For instance, classical feedback controllers for noisy systems or also MPC for stochastic systems based on the Koopman generator (cf. “Klus et al. Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. Physica D, 2020.”) • I believe that it would be very interesting to see the resulting optimal control signal $u = \Psi(x) * w/v$, to give readers a better intuition how the controller performs and how regular this solution looks. • The language quality is insufficient in many many places and should be improved significantly. A few examples (by far not all!) are: ◦ Weineer Process → Wiener Process ◦ The references to previous Theorems are very inconsistent (e.g., ([20], Theorem 21.11)). Please use the latex command to unify this ([?, Theorem 21.11]. ◦ “In particular, the vector fields $f(x)$ and $n(x)$ are assumed to satisfy the following Lipschitz condition” ◦ The term RBF is never introduced ◦ “The number of basis functions along each dimension is chosen to be 15×15 for 2D example”. It's 15 in each dimension and not 15×15 ! Limited circulation. Attachment to ACC 2021 Review 2083 Limited circulation