Covariance Steering for Nonlinear Control-affine Systems

Hongzhe Yu, Zhenyang Chen, and Yongxin Chen

Abstract—We consider the covariance steering problem for nonlinear control-affine systems. Our objective is to find an optimal control strategy to steer the state of a system from an initial distribution to a target one whose mean and covariance are given. Due to the nonlinearity, the existing techniques for linear covariance steering problems are not directly applicable. By leveraging the Girsanov theorem, we formulate the problem into an optimization over the space of path distributions. We then adopt a generalized proximal gradient algorithm to solve this optimization, where each update requires solving a linear covariance steering problem. Our algorithm is guaranteed to converge to a local solution with a sublinear rate. In addition, each iteration of the algorithm can be achieved in closed form, and thus the computational complexity of it is insensitive to the resolution of time-discretization. In the examples, our method achieves 1000 times speedup over an existing algorithm.

I. Introduction

Uncertainties are ubiquitous in engineering systems and an important task of feedback control is to mitigate their effects on the behavior of the systems. Very often the goal of control is to ensure the system does not deviate too much from the desired behavior in the presence of uncertainties. A standard approach for this task is optimal control where a cost function is used to promote the desired system behavior. The cost function is a design parameter and is often chosen by trial and error until the desired performance is achieved. The covariance control/assignment paradigm was born [1], [2], [3], [4] in the quest of providing a more direct approach to mitigate uncertainties. In its original formulation, the goal was to find a control strategy for a linear time-invariant stochastic system so that its state has a specified covariance in steady state. The covariance control theory was recently extended to the finite horizon control setting [5], [6], [7], [8], [9], [10], [11] where one seeks to steer the state covariance of a linear dynamic system from an initial value to a target value. The covariance control framework has been successfully used in a range of applications [12], [13], [14], [15].

Most existing works on covariance control/steering are for linear dynamics. The methods developed in them are not directly applicable to nonlinear problems. The purpose of this work is to develop a covariance steering method for nonlinear stochastic dynamics. Even though for nonlinear stochastic systems the state distribution is no longer Gaussian, the covariance still provides valuable information about the uncertainties of the state variables. Indeed, under the mild assumption that the state follows a sub-Gaussian distribution, the covariance can properly capture the high probability region and thus the uncertainty of the state [16]. The same idea of approximating the full distribution with first order and second order moments has been widely used in filtering algorithms such as the extended Kalman filter algorithm.

In this paper, we develop an efficient algorithm of nonlinear covariance steering for control-affine nonlinear systems. Our method is based on the celebrated Girsanov theorem [17] for stochastic processes, which connects the control energy with the Kullback-Leibler divergence [18] between the measures induced by the con-

trolled and uncontrolled dynamics. As a result, the covariance steering problem we consider can be reformulated as an infinite dimensional optimization problem over the space of measures of trajectories. The optimization can then be solved using the generalized proximal gradient descent algorithm [19], [20], a popular algorithm for nonlinear programming. For our covariance steering problem, the proximal gradient descent algorithm converts the problem into a sequence of linear covariance steering problems, and each of these linear covariance steering problems has a closed-form solution. Our algorithm is guaranteed to converge to a local optimal solution with sublinear rate $\mathcal{O}(1/k)$. Moreover, each iteration of the algorithm is associated with a feasible solution to the problem. Thus, whenever we stop the algorithm, we can still obtain a controller that steer the state to desired target mean and covariance.

The nonlinear covariance control problem was recently considered in [21], [22], [23], [24]. In [21] the cost function is assumed to be quadratic and an iterative linearization method resembling sequential quadratic programing is used. The algorithm is not guaranteed to converge. In [22] a method based on dynamic differential programming was developed. To enforce the covariance constraint on the terminal state, the algorithm first relaxes the constraint using Lagrangian method and then uses primal-dual updates to search for the solution. In [23], nonlinear covariance steering for discrete dynamics was studied. The method is similar to that in [21] but an unscented transform is used to better capture the nonlinearity of the dynamics. The method was extended to the data-driven setting in [24] where a Gaussian process is used to identify the unknown dynamics. Compared with all these works, our method is tailored for continuous-time control-affine dynamics with general state cost. A distinguishing feature of the proposed algorithm is that its complexity is insensitive to the resolution of time-discretization in implementation while all the other algorithms do not scale well with this resolution. In our experiments, our method shows at least 1000 times acceleration compared with [21].

The rest of the paper is structured as follows. In Section II we provide the background on covariance control for linear dynamics and the proximal gradient algorithm. The problem is formulated in Section III and the algorithm is developed in Section IV. An extension of the method to a slightly more general systems is presented in Section V, which is followed by a numerical example in Section VI and a concluding remark in Section VII.

II. Background

In this section, we provide a brief introduction to the covariance steering problem for continuous-time linear systems [5], [6], [7]. We also present key steps of the proximal gradient algorithm, on which our method is based.

A. Covariance steering for linear systems

In the covariance steering/control problems for linear systems [5], the objective is to drive the state of a linear stochastic system from an initial Gaussian random vector $X_0 \sim \mathcal{N}(m_0, \Sigma_0)$ at $t = 0$

to a terminal one $X_1 \sim \mathcal{N}(m_1, \Sigma_1)$ at $t = 1$ over a time interval $[0, 1]$ [1]. In [5], [7], the linear dynamics under consideration is

$$dX_t = A(t)X_t dt + B(t)(u_t dt + \sqrt{\epsilon} dW_t),$$

which describes the behavior of linear stochastic control system whose actuation is corrupted by white noise. Here the pair $A(t) \in \mathbb{R}^{n \times n}, B(t) \in \mathbb{R}^{n \times p}$ is assumed to be controllable, $W_t \in \mathbb{R}^p$ represents a standard Wiener process [25], and $\epsilon > 0$ parameterizes the intensity of the disturbance. When one seeks also to minimize a certain cost function, the covariance steering problem for linear dynamics in the continuous-time setting is [7]

$$\min_u \quad \mathbb{E}\left\{ \int_0^1 [\frac{1}{2}\|u_t\|^2 + \frac{1}{2}X_t^T Q(t)X_t]dt \right\} \tag{1a}$$
$$dX_t = A(t)X_t dt + B(t)(u_t dt + \sqrt{\epsilon} dW_t) \tag{1b}$$
$$X_0 \sim \mathcal{N}(m_0, \Sigma_0), \quad X_1 \sim \mathcal{N}(m_1, \Sigma_1), \tag{1c}$$

where the optimization is over all the state feedback control strategies. The state cost matrix $Q(t) \in \mathbb{R}^{n \times n}$ is often assumed to be positive semi-definite. This problem differs from standard linear quadratic optimal control problems in that there is no explicit terminal cost. Instead, a constraint $X_1 \sim \mathcal{N}(m_1, \Sigma_1)$ on the statistics of the terminal state is imposed.

Thanks to linearity, the mean and covariance of the system can be controlled separately [5], [6], [7]. The control of the mean is the solution to the standard optimal control problem

$$\min_v \quad \int_0^1 [\frac{1}{2}\|v_t\|^2 + \frac{1}{2}x_t^T Q(t)x_t]dt \tag{2a}$$
$$\dot{x}_t = A(t)x_t + B(t)v_t \tag{2b}$$
$$x_0 = m_0, \quad x_1 = m_1. \tag{2c}$$

The control for the covariance is of the form $-B(t)^T \Pi(t)X_t$ with $\Pi(\cdot), \mathrm{H}(\cdot)$ satisfying a coupled Riccati equations [9]

$$-\dot{\Pi}(t) = A^T \Pi(t) + \Pi(t)A(t) - \Pi(t)B(t)B(t)^T \Pi(t) + Q(t) \tag{3a}$$
$$-\dot{\mathrm{H}}(t) = A^T \mathrm{H}(t) + \mathrm{H}(t)A(t) + \mathrm{H}(t)B(t)B(t)^T \mathrm{H}(t) - Q(t) \tag{3b}$$
$$\epsilon \Sigma_0^{-1} = \Pi(0) + \mathrm{H}(0) \tag{3c}$$
$$\epsilon \Sigma_1^{-1} = \Pi(1) + \mathrm{H}(1). \tag{3d}$$

It turns out (3) has a unique solution as in the following result [7].

Theorem 1. *The coupled system of Riccati equations (3a-3d) has a unique solution, which is determined by the initial value problem consisting of (3a-3b) and*

$$\Pi(0) = \frac{\epsilon \Sigma_0^{-1}}{2} - \Phi_{12}^{-1} \Phi_{11} \tag{4a}$$
$$\qquad - \Sigma_0^{-1/2} \left(\frac{\epsilon^2 I}{4} + \Sigma_0^{1/2} \Phi_{12}^{-1} \Sigma_1 (\Phi_{12}^T)^{-1} \Sigma_0^{1/2} \right)^{1/2} \Sigma_0^{-1/2},$$
$$\mathrm{H}(0) = \epsilon \Sigma_0^{-1} - \Pi(0), \tag{4b}$$

where

$$\Phi(t, s) = \begin{bmatrix} \Phi_{11}(t, s) & \Phi_{12}(t, s) \\ \Phi_{21}(t, s) & \Phi_{22}(t, s) \end{bmatrix}$$

is the state transition matrix corresponding to $\partial \Phi(t, s)/\partial t = M(t)\Phi(t, s)$ *with* $\Phi(s, s) = I$ *and*

$$M(t) = \begin{bmatrix} A(t) & -B(t)B(t)^T \\ -Q(t) & -A(t)^T \end{bmatrix},$$

[1] Any finite time interval $[0, T]$ can be converted to $[0, 1]$ by rescaling. Thus, without loss of generality, we use the time interval $[0, 1]$ throughout.

and

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} := \begin{bmatrix} \Phi_{11}(1, 0) & \Phi_{12}(1, 0) \\ \Phi_{21}(1, 0) & \Phi_{22}(1, 0) \end{bmatrix}.$$

Denote the optimal control for mean to be $v_t^\star$ and corresponding state trajectory to be $x_t^\star$. Combining it with the covariance control yields the optimal feedback policy

$$u_t^\star = -B(t)^T \Pi(t)(X_t - x_t^\star) + v_t^\star.$$

B. Proximal gradient algorithm

The proximal gradient algorithm [20] is designed for the composite optimization

$$\min_{y \in \mathcal{Y}} F(y) + G(y), \tag{5}$$

where $\mathcal{Y}$ denotes a feasible set. The function F is assumed to be smooth. The function G is usually a regularizer that is often not smooth. The algorithm follows the update

$$y^{k+1} = \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{2\eta}\|y - (y^k - \eta \nabla F(y^k))\|^2 \tag{6a}$$
$$= \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{2\eta}\|y - y^k\|^2 + \langle \nabla F(y^k), y - y^k \rangle \tag{6b}$$

with $\eta > 0$ being the stepsize. One advantage of the proximal gradient algorithm is that it only evaluates the gradient of F and doesn't require even the differentiability of G. In many applications, G is a regularizer of simple form, e.g., 1-norm, nuclear norm, and the minimization (6) can be carried out efficiently.

The proximal gradient algorithm has been generalized to the non-Euclidean setting. It is built upon the mirror descent method [19], [20]. Let $D(\cdot, \cdot)$ be a Bregman divergence, then the generalized non-Euclidean proximal gradient algorithm reads

$$y^{k+1} = \operatorname{argmin}_{y \in \mathcal{Y}} G(y) + \frac{1}{\eta}D(y, y^k) + \langle \nabla F(y^k), y - y^k \rangle. \tag{7}$$

A popular choice of $D(\cdot, \cdot)$ is the Kullback-Leibler divergence $\mathrm{KL}(\cdot \| \cdot)$, which is suitable for optimization over probability vectors/distributions.

The proximal gradient algorithm enjoys nice convergence properties. When both F and G are convex, the algorithm is guaranteed to converge to the global minimum with rate $\mathcal{O}(1/k)$ [19], [20]. When F is nonconvex, one can only expect for convergence to local solutions. It turns out that objective function $F(y) + G(y)$ is monotonically decreasing along the updates, and the updates converge to some stationary points with sublinear rate $\mathcal{O}(1/k)$ with respect to some suitable criteria [26].

III. PROBLEM FORMULATION

Consider a nonlinear control-affine system

$$dX_t = f(t, X_t)dt + B(t)(u_t dt + \sqrt{\epsilon} dW_t) \tag{8}$$

where $X_t \in \mathbb{R}^n$ represents the state, and W_t denotes a standard Wiener process. The drift function $f(t, x)$ is assumed to be locally Lipschitz continuous with respect to x and continuous with respect to t. The input matrix $B(t) \in \mathbb{R}^{n \times p}$ is assumed to be continuous and full rank. Many nonlinear dynamics in engineering applications can be modeled by (8) where $\sqrt{\epsilon} dW_t$ captures the uncertainties in the actuation. Note that the noise is assumed to enter the dynamics in the same channel as the input. This is a widely used assumption in stochastic control in both theory and practice [27], [28], [29], [30].

We are interested in the covariance control problems of steering the state statistics of the system from an initial value to a target one. In particular, we seek a state feedback control strategy that achieves this goal and meanwhile minimizes a cost function. The covariance steering problem we consider is

$$\min_{u} \quad \mathbb{E}\left\{\int_0^1 [\frac{1}{2}\|u_t\|^2 + V(X_t)]dt\right\} \tag{9a}$$

$$dX_t = f(t, X_t)dt + B(t)(u_t dt + \sqrt{\epsilon}dW_t) \tag{9b}$$

$$X_0 \sim \rho_0, \quad X_1 \sim \rho_1, \tag{9c}$$

where ρ_0 (ρ_1) is a probability distribution with mean m_0 (m_1) and covariance Σ_0 (Σ_1). The cost function has two decoupled terms: $\frac{1}{2}\|u_t\|^2$ that only depends on the control and $V(X_t)$ that only depends on the state. The problem (9) has been investigated in the study of distribution control [27], [31], [32], [30] for general distributions ρ_0, ρ_1, which links control theory and optimal transport theory [27], [33].

The special structure of the cost and dynamics in (9) leads to an elegant reformulation as an optimization over probability measures. Denote by $\mathcal{P}^u$ the distribution over the path space $\Omega := C([0, 1], \mathbb{R}^n)$ induced by the stochastic process (8), and as a convention, by $\mathcal{P}^0$ the distribution induced by the same process with zero control, then the Girsanov theorem [17], [34] states that

$$\frac{d\mathcal{P}^u}{d\mathcal{P}^0} = \exp\left(\int_0^1 \frac{1}{2\epsilon}\|u_t\|^2 dt + \frac{1}{\sqrt{\epsilon}}u_t^T dW_t\right). \tag{10a}$$

When B is full (column) rank, this can be equivalently written as

$$\frac{d\mathcal{P}^u}{d\mathcal{P}^0} = \exp\left(\int_0^1 \frac{1}{2\epsilon}\|Bu_t\|^2_{(BB^T)^\dagger}dt + \frac{1}{\sqrt{\epsilon}}u_t^T dW_t\right). \tag{10b}$$

It follows from the Girsanov theorem that [35], [36], [27]

$$\mathbb{E}\left\{\int_0^1 \frac{1}{2\epsilon}\|u_t\|^2 dt\right\} = \mathrm{KL}(\mathcal{P}^u\|\mathcal{P}^0) := \int \log \frac{d\mathcal{P}^u}{d\mathcal{P}^0}d\mathcal{P}^u. \tag{11}$$

Intuitively, it says that the difference between the controlled and uncontrolled processes can be quantified by the control energy. This relation is heavily utilized recently in the distributional control problem [27]. Thus, the problem (9) can be reformulated as

$$\min_{\mathcal{P}^u} \quad \int d\mathcal{P}^u \left[\log \frac{d\mathcal{P}^u}{d\mathcal{P}^0} + \frac{1}{\epsilon}V\right] \tag{12a}$$

$$(X_0)_\sharp \mathcal{P}^u = \rho_0, \quad (X_1)_\sharp \mathcal{P}^u = \rho_1. \tag{12b}$$

Here $(X_0)_\sharp \mathcal{P}^u$ ($(X_1)_\sharp \mathcal{P}^u$) stands for the distribution of X_0 (X_1) when the process X_t is associated with the distribution $\mathcal{P}^u$.

Denote by $\Pi(\rho_0, \rho_1)$ the set of all distributions over the path space Ω such that the constraints (12b) hold. Let

$$F(\mathcal{P}^u) = \int [\frac{1}{\epsilon}V - \log d\mathcal{P}^0]d\mathcal{P}^u$$

and

$$G(\mathcal{P}^u) = \int d\mathcal{P}^u \log d\mathcal{P}^u. \tag{13}$$

Then, (12) becomes a composite optimization

$$\min_{\mathcal{P}^u \in \Pi(\rho_0, \rho_1)} F(\mathcal{P}^u) + G(\mathcal{P}^u). \tag{14}$$

This is an infinite dimensional optimization. For general marginals ρ_0, ρ_1 and nonlinear dynamics (8), the problem can be addressed using the distribution control framework [27], [30]. However, this method requires solving a coupled partial differential equation system and does not scale to cases with large state dimension n [27], [30].

Our goal is to develop an efficient algorithm to approximately solve (14). In the covariance steering formulation, instead of solving (14) exactly, we look for an approximate solution $\mathcal{P}^u$ that is induced by a Gaussian Markov process. Thus, we restrict our search space to $\hat{\Pi}(\rho_0, \rho_1)$, the space of measures over the path space Ω that are induced by Gaussian Markov processes and have marginal distributions $\mathcal{N}(m_0, \Sigma_0), \mathcal{N}(m_1, \Sigma_1)$. When $\mathcal{P}$ has relatively small variance, $F(\mathcal{P})$ can be approximated by

$$\int [\frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0]d\mathcal{P}$$

where $\hat{V}$ is the second order approximation of V and $\hat{\mathcal{P}}^0$ is a Gaussian Markov approximation of $\mathcal{P}^0$, both along the mean of $\mathcal{P}$. More precisely, let the mean of $\mathcal{P}$ be the trajectory z_t, then

$$\hat{V}(t, x) = V(z_t) + (x^T - z_t^T)\nabla V(z_t) + \frac{1}{2}(x^T - z_t^T)\nabla^2 V(z_t)(x - z_t), \tag{15}$$

and $\hat{\mathcal{P}}^0$ is associated with the Gaussian Markov process

$$dX_t = \nabla f(t, z_t)^T X_t dt + [f(t, z_t) - \nabla f(t, z_t)^T z_t]dt + \sqrt{\epsilon}B(t)dW_t. \tag{16}$$

We have thus arrived at the following nonlinear covariance steering problem

$$\min_{\mathcal{P}^u \in \hat{\Pi}(\rho_0, \rho_1)} \int [\frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0]d\mathcal{P}^u + \int d\mathcal{P}^u \log d\mathcal{P}^u. \tag{17}$$

This nonlinear covariance steering problem turns out to be considerably easier than the distribution control problem (14) from a computational point of view. Indeed, in the following section, we develop an efficient algorithm for it that is scalable to problems with large state dimension.

IV. PROXIMAL GRADIENT ALGORITHM FOR COVARIANCE STEERING

The nonlinear covariance steering problem (17) is clearly a composite optimization

$$\min_{\mathcal{P}^u \in \hat{\Pi}(\rho_0, \rho_1)} F(\mathcal{P}^u) + G(\mathcal{P}^u) \tag{18}$$

where G is as in (13), and by abuse of notation,

$$F(\mathcal{P}^u) = \int [\frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0]d\mathcal{P}^u = \langle \frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0, \mathcal{P}^u\rangle. \tag{19}$$

In this section, we develop an efficient algorithm to solve (18) based on the generalized proximal gradient algorithm (7).

A. Main algorithm

Since (18) is an optimization over the space of probability measures, we use Kullback-Leibler divergence in (7). This leads to the following iteration

$$\mathcal{P}_{k+1} = \mathrm{argmin}_{\mathcal{P} \in \hat{\Pi}(\rho_0, \rho_1)} G(\mathcal{P}) + \frac{1}{\eta}\mathrm{KL}(\mathcal{P}\|\mathcal{P}_k) + \langle \frac{\delta F}{\delta \mathcal{P}}(\mathcal{P}_k), \mathcal{P}\rangle. \tag{20}$$

Here we use variation $\frac{\delta F}{\delta \mathcal{P}}$ instead of gradient since $\mathcal{P}$ is an infinite dimensional object. As we will see below, each iteration can be realized by solving a linear covariance steering problem.

First, we derive an explicit expression for $\frac{\delta F}{\delta \mathcal{P}}$. Note that both $\hat{V}$ and $\hat{\mathcal{P}}^0$ depend on the mean of $\mathcal{P}$, thus $F(\mathcal{P})$ is not a linear function of $\mathcal{P}$. The variation $\frac{\delta F}{\delta \mathcal{P}}$ thus has some extra terms other than $\frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0$.

Lemma 2. *Let $\mathcal{P}$ be the path distribution induced by a Gaussian Markov process*

$$dX_t = A(t)X_t dt + a(t)dt + \sqrt{\epsilon}B(t)dW_t,$$

then the variation of $F(\mathcal{P})$ with respect to $\mathcal{P}$ is

$$\frac{\delta F}{\delta \mathcal{P}}(\mathcal{P})(t,x) = \frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0 + \frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}(\nabla^2 V(z_t)\Sigma_t) \quad (21)$$
$$+ \frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}((BB^T)^\dagger(\nabla f(z_t)^T - A)\Sigma_t(\nabla f(z_t) - A^T)),$$

where $(\cdot)^\dagger$ represents pseudo-inverse, z_t and Σ_t are the mean and covariance of $\mathcal{P}$, respectively, and $\hat{V}, \hat{\mathcal{P}}^0$ are as in (15) and (16).

We remark that the extra terms involve the second order derivative of the drift f and the third order derivative of V. Thus, if the dynamics is linear and the cost is quadratic as in the linear covariance steering problems, $\frac{\delta F}{\delta \mathcal{P}}(\mathcal{P}) = \frac{1}{\epsilon}\hat{V} - \log d\hat{\mathcal{P}}^0$ and is independent of $\mathcal{P}$. We also remark that $\frac{\delta F}{\delta \mathcal{P}}$ in (21) is a quadratic function.

We next establish that each iteration of (20) is a covariance control problem for linear dynamics. To this end, we represent $\mathcal{P}_k^u$ by a Gaussian Markov process

$$dX_t = A_k(t)X_t dt + a_k(t)dt + \sqrt{\epsilon}B(t)dW_t. \quad (22)$$

Denote its mean and covariance by z^k and Σ^k respectively, then

$$\dot{z}_t^k = A_k(t)z_t^k + a_k(t) \quad (23a)$$
$$\dot{\Sigma}_t^k = A_k(t)\Sigma_t^k + \Sigma_t^k A_k(t)^T + \epsilon B(t)B(t)^T. \quad (23b)$$

Plugging (21) into (20), in view of representation (22) of $\mathcal{P}_k^u$, we obtain

$$\mathcal{P}_{k+1}^u = \mathrm{argmin}_{\mathcal{P}\in\hat{\Pi}(\rho_0,\rho_1)}\int[\frac{1}{\epsilon}\hat{V} + \frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}(\nabla^2 V(z_t^k)\Sigma_t^k)$$
$$+ \frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}((BB^T)^\dagger(\nabla f(z_t^k)^T - A_k)\Sigma_t^k(\nabla f(z_t^k) - A_k^T))$$
$$- \log d\hat{\mathcal{P}}^0 - \frac{1}{\eta}\log d\mathcal{P}_k^u]d\mathcal{P} + (1+\frac{1}{\eta})\int d\mathcal{P}\log d\mathcal{P}. \quad (24)$$

Linearizing the uncontrolled system

$$dX_t = f(t,X_t)dt + \sqrt{\epsilon}B(t)dW_t \quad (25)$$

along z^k yields the linear approximation

$$dX_t = \hat{A}_k(t)X_t dt + \hat{a}_k(t)dt + \sqrt{\epsilon}B(t)dW_t \quad (26)$$

with

$$\hat{A}_k(t) = \nabla f(t,z_t^k)^T \quad (27a)$$
$$\hat{a}_k(t) = f(t,z_t^k) - \nabla f(t,z_t^k)^T z_t^k. \quad (27b)$$

This linear approximation corresponds to $\hat{\mathcal{P}}^0$ at the k-th iteration. This is the same as (16) with some additional notations being introduced to emphasize the dependency on the iteration number k. With these Gaussian Markov representations of $\mathcal{P}^0, \mathcal{P}_k^u$ we establish the following.

Theorem 3. *Each proximal gradient iteration (24) amounts to solving the following linear covariance steering problem*

$$\min_u \quad \mathbb{E}\left\{\int_0^1[\frac{1}{2}\|u_t\|^2 + \frac{1}{2}X_t^T Q_k(t)X_t + X_t^T r_k(t)]dt\right\} \quad (28a)$$
$$dX_t = \frac{1}{1+\eta}[A_k(t) + \eta\hat{A}_k(t)]X_t dt + \frac{1}{1+\eta}[a_k(t) \quad (28b)$$
$$+ \eta\hat{a}_k(t)]dt + B(t)(u_t dt + \sqrt{\epsilon}dW_t)$$
$$X_0 \sim \rho_0, \quad X_1 \sim \rho_1, \quad (28c)$$

where

$$Q_k(t) = \frac{\eta}{(1+\eta)}\nabla^2 V(z_t^k) +$$
$$\frac{\eta}{(1+\eta)^2}(A_k(t) - \hat{A}_k(t))^T(B(t)B(t)^T)^\dagger(A_k(t) - \hat{A}_k(t))$$

and

$$r_k(t) = \frac{\eta}{1+\eta}\nabla V + \frac{\eta}{2(1+\eta)}[\nabla\mathrm{Tr}(\nabla^2 V\Sigma_t^k) - 2\nabla^2 V(z_t^k)z_t^k$$
$$+ \nabla\mathrm{Tr}((B(t)B(t)^T)^\dagger(\nabla f^T - A_k)\Sigma_t^k(\nabla f - A_k^T))]$$
$$+ \frac{\eta}{(1+\eta)^2}(A_k(t) - \hat{A}_k(t))^T(B(t)B(t)^T)^\dagger(a_k(t) - \hat{a}_k(t)).$$

Proof. The optimization in (24) is equivalent to

$$\min_{\mathcal{P}\in\hat{\Pi}(\rho_0,\rho_1)}\int[\frac{1}{\epsilon}\hat{V} + \frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}(\nabla^2 V(z_t^k)\Sigma_t^k) +$$
$$\frac{1}{2\epsilon}x^T\nabla\mathrm{Tr}((BB^T)^\dagger(\nabla f(z_t^k)^T - A_k)\Sigma_t^k(\nabla f(z_t^k) - A_k^T))]d\mathcal{P}$$
$$+ \mathrm{KL}(\mathcal{P}\|\hat{\mathcal{P}}^0) + \frac{1}{\eta}\mathrm{KL}(\mathcal{P}\|\mathcal{P}_k^u). \quad (29)$$

Let $\mathcal{P} = \mathcal{P}^u$ be parametrized by (28b), then, by Girsanov theorem (10), $\mathrm{KL}(\mathcal{P}^u\|\hat{\mathcal{P}}^0)$ equals

$$\frac{1}{2\epsilon}\mathbb{E}\left\{\int_0^1\|\frac{1}{1+\eta}(A_kX_t - \hat{A}_kX_t + a_k - \hat{a}_k) + Bu_t\|^2_{(BB^T)^\dagger}dt\right\} \quad (30)$$

and $\mathrm{KL}(\mathcal{P}^u\|\mathcal{P}_k^u)$ equals

$$\frac{1}{2\epsilon}\mathbb{E}\left\{\int_0^1\|\frac{\eta}{1+\eta}(\hat{A}_kX_t - A_kX_t + \hat{a}_k - a_k) + Bu_t\|^2_{(BB^T)^\dagger}dt\right\} \quad (31)$$

Combining (30) and (31) points to

$$\mathrm{KL}(\mathcal{P}^u\|\hat{\mathcal{P}}^0) + \frac{1}{\eta}\mathrm{KL}(\mathcal{P}^u\|\mathcal{P}_k^u)$$
$$= \frac{1}{2\epsilon}\mathbb{E}\left\{\int_0^1[(1+\frac{1}{\eta})\|u_t\|^2 +$$
$$\frac{1}{1+\eta}\|(A_kX_t - \hat{A}_kX_t + a_k - \hat{a}_k)\|^2_{(B(t)B(t)^T)^\dagger}]dt\right\}.$$

In view of (15), plugging the above into (29) yields the objective function (28a), after removing some constant terms and rescaling with ratio $(1+\eta)/(\epsilon\eta)$. $\square$

The linear covariance steering problem (28) has a closed-form solution (see Section IV-B for more details). Denote the optimal control policy to (28) by

$$u_t^\star = K_k(t)X_t + d_k(t), \quad (32)$$

then $\mathcal{P}_{k+1}^u$ is induced by the closed-loop process

$$dX_t = \frac{1}{1+\eta}[A_k(t) + \eta\hat{A}_k(t)]X_t dt + \frac{1}{1+\eta}[a_k(t) + \eta\hat{a}_k(t)]dt$$
$$+ B(t)(u_t^\star dt + \sqrt{\epsilon}dW_t).$$

It follows that

$$A_{k+1}(t) = \frac{1}{1+\eta}[A_k(t) + \eta\hat{A}_k(t)] + B(t)K_k(t) \quad (33a)$$
$$a_{k+1}(t) = \frac{1}{1+\eta}[a_k(t) + \eta\hat{a}_k(t)] + B(t)d_k(t). \quad (33b)$$

The above equations (33) provide an extremely simple rule to update $\mathcal{P}_k^u$, based on which we present the proximal gradient algorithm (Algorithm 1) for nonlinear covariance steering.

There are many choices for initialization. For instance, one can set $\mathcal{P}_0^u$ to be the process $dX_t = \sqrt{\epsilon}B(t)dW_t$, which means $A_0(t) \equiv 0$ and $a_0(t) \equiv 0$. An alternative option is a linearization of the prior process (25). More precisely, set z_t^0 to be the solution to $\dot{z}_t^0 = f(t, z_t^0)$ and $A_0(t) = \nabla f(t, z_t^0)^T$. The initial $a_0(t)$ can be calculated by

$$a_0(t) = f(t, z_t^0) - A_0(t)z_t^0.$$

Algorithm 1 Proximal gradient nonlinear covariance steering algorithm

---

    Initialize A_0, a_0
    for $k = 1, 2, \ldots$ **do**
        Update A_k using (33a)
        Update a_k using (33b)
    end for

Remark 4. *Algorithm 1 inherits the convergence properties of the proximal gradient algorithm and thus converges to a local solution to (17) with sublinear rate $\mathcal{O}(1/k)$.*

Remark 5. *The computational cost of Algorithm 1 can be divided into two parts: i) linearization to get $\hat{A}_k, \hat{a}_k$ and ii) obtaining optimal strategy (32) for the linear covariance steering problem. Both have closed-form and can be computed efficiently, even for high resolution of time-discretization. This is a major advantage over existing works on nonlinear covariance steering.*

B. Covariance steering for linear systems with drift

The linear covariance control problem in Theorem 3 is slightly more general than (1) due to some extra terms in the dynamics and cost. To provide a closed-form solution to (28), we next extend the results in Section II-A to the following variation of linear covariance steering problems with a drift term in the dynamics and an extra linear term in the cost

$$\min_u \; \mathbb{E}\left\{\int_0^1 [\frac{1}{2}\|u_t\|^2 + \frac{1}{2}X_t^T Q(t)X_t + X_t^T r(t)]dt\right\} \quad (34a)$$
$$dX_t = A(t)X_t dt + a(t)dt + B(t)(u_t dt + \sqrt{\epsilon}dW_t) \quad (34b)$$
$$X_0 \sim \mathcal{N}(m_0, \Sigma_0), \quad X_1 \sim \mathcal{N}(m_1, \Sigma_1). \quad (34c)$$

By linearity, the mean and the covariance can be controlled separately. Apparently, the covariance control part is the same as Problem (1) since both the dynamics of covariance and the cost related to covariance in (34) are exactly the same as (1).

The deterministic control for the mean is slightly different from (1). Let v_t be the mean of the control, then it is associated with the deterministic optimal control problem

$$\min_v \quad \int_0^1 [\frac{1}{2}\|v_t\|^2 + \frac{1}{2}x_t^T Q(t)x_t + x_t^T r(t)]dt \quad (35a)$$
$$\dot{x}_t = A(t)x_t + a(t) + B(t)v_t \quad (35b)$$
$$x_0 = m_0, \quad x_1 = m_1. \quad (35c)$$

The above problem can be solved using the Pontryagin's principle. More specifically, it amounts to solving the differential equations

$$\begin{bmatrix} \dot{x}_t \\ \dot{\lambda}_t \end{bmatrix} = \begin{bmatrix} A(t) & -B(t)B(t)^T \\ -Q(t) & -A(t)^T \end{bmatrix} \begin{bmatrix} x_t \\ \lambda_t \end{bmatrix} + \begin{bmatrix} a(t) \\ -r(t) \end{bmatrix} \quad (36)$$

with boundary condition $x_0 = m_0, x_1 = m_1$. Using the notation in Section II we obtain

$$\begin{bmatrix} x_1 \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} x_0 \\ \lambda_0 \end{bmatrix} + \int_0^1 \begin{bmatrix} \Phi_{11}(1,\tau) & \Phi_{12}(1,\tau) \\ \Phi_{21}(1,\tau) & \Phi_{22}(1,\tau) \end{bmatrix} \begin{bmatrix} a(\tau) \\ -r(\tau) \end{bmatrix} d\tau.$$

It follows

$$\lambda_0 = \Phi_{12}^{-1}\left(m_1 - \Phi_{11}m_0 - \int_0^1(\Phi_{11}(1,\tau)a(\tau) - \Phi_{12}(1,\tau)r(\tau))d\tau\right).$$

Plugging it back to (36) gives the expression for λ_t. The solution to the deterministic control problem (35) is $v_t^\star = -B(t)^T\lambda_t$. The optimal control strategy for the covariance steering problem (34) is

$$u_t^\star = -B(t)^T\Pi(t)(X_t - x_t^\star) + v_t^\star.$$

C. Construct optimal control

Upon convergence, the total feedback control policy can be recovered as follows. Denote $A^\star, a^\star$ the limit point of $\{A_k, a_k\}$. Let $z_t^\star, \Sigma_t^\star$ be the solutions to

$$\dot{z}_t^\star = A^\star(t)z_t^\star + a^\star(t),$$
$$\dot{\Sigma}_t^\star = A^\star(t)\Sigma_t^\star + \Sigma_t^\star A^\star(t)^T + \epsilon B(t)B(t)^T,$$

then they satisfy that $z_0^\star = m_0, z_1^\star = m_1$ and $\Sigma_0^\star = \Sigma_0, \Sigma_1^\star = \Sigma_1$. Denote the linearization of the prior dynamics (25) along $z_t^\star$ as

$$dX_t = \hat{A}^\star(t)X_t dt + \hat{a}^\star(t)dt + \sqrt{\epsilon}B(t)dW_t, \quad (37)$$

then

$$\hat{A}^\star(t) = \nabla f(t, z_t^\star)^T \quad (38a)$$
$$\hat{a}^\star(t) = f(t, z_t^\star) - \nabla f(t, z_t^\star)^T z_t^\star. \quad (38b)$$

The distribution $\mathcal{P}^\star$ over the path space Ω associated with

$$dX_t = A^\star(t)X_t dt + a^\star(t)dt + \sqrt{\epsilon}B(t)dW_t \quad (39)$$

is the solution to our covariance steering problem (17). To retrieve the optimal control strategy, consider the optimization

$$\min_{\mathcal{P}\in\hat{\Pi}(\rho_0,\rho_1)} G(\mathcal{P}) + \langle\frac{\delta F}{\delta \mathcal{P}}(\mathcal{P}^\star), \mathcal{P}\rangle. \quad (40)$$

It shares the same minimizer, which is $\mathcal{P}^\star$, as

$$\min_{\mathcal{P}\in\hat{\Pi}(\rho_0,\rho_1)} G(\mathcal{P}) + \langle\frac{\delta F}{\delta \mathcal{P}}(\mathcal{P}^\star), \mathcal{P}\rangle + \frac{1}{\eta}\mathrm{KL}(\mathcal{P}\|\mathcal{P}^\star). \quad (41)$$

Indeed, since $\mathcal{P}^\star$ is the minimizer of (41) and $\mathrm{KL}(\mathcal{P}\|\mathcal{P}^\star)$ vanishes when $\mathcal{P} = \mathcal{P}^\star$, $\mathcal{P}^\star$ must be a minimizer of (40) as well. To retrieve the final optimal control from (40), we reformulate it as a linear covariance control problem as follows. The proof is similar to that of Theorem 3 and is thus omitted.

Proposition 1. *The optimization (40) amounts to the linear covariance steering problem*

$$\min_u \; \mathbb{E}\left\{\int_0^1 [\frac{1}{2}\|u_t\|^2 + \frac{1}{2}X_t^T Q^\star(t)X_t + X_t^T r^\star(t)]dt\right\} \quad (42a)$$
$$dX_t = \hat{A}^\star(t)X_t dt + \hat{a}^\star(t)dt + B(t)(u_t dt + \sqrt{\epsilon}dW_t) \quad (42b)$$
$$X_0 \sim \rho_0, \quad X_1 \sim \rho_1, \quad (42c)$$

where $Q^\star(t) = \nabla^2 V(z_t^\star)$ and

$$r^\star(t) = \nabla V(z_t^\star) + \frac{1}{2}[\nabla \mathrm{Tr}(\nabla^2 V(z_t^\star)\Sigma_t^\star) - 2\nabla^2 V(z_t^\star)z_t^\star$$
$$+ \nabla \mathrm{Tr}((B(t)B(t)^T)^\dagger(\nabla f^T - A^\star)\Sigma_t^\star(\nabla f - A^{\star T}))].$$

Denote by

$$u_t^\star = K^\star(t)X_t + d^\star(t) \quad (43)$$

the optimal control for the linear covariance control problem in Proposition 1, then it is also the optimal control strategy to the nonlinear covariance steering problem (17).

V. GENERALIZATIONS

The nonlinear covariance steering algorithm can be extended to the general control-affine system

$$dX_t = f(t, X_t)dt + g(t, x)(u_t dt + \sqrt{\epsilon}dW_t). \qquad (44)$$

The goal is again to approximately solve the density control problem

$$\min_{u_t} \quad \mathbb{E}\left\{\int_0^1 [\frac{1}{2}\|u_t\|^2 + V(X_t)]dt\right\} \qquad (45a)$$

$$dX_t = f(t, X_t)dt + g(t, X_t)(u_t dt + \sqrt{\epsilon}dW_t) \quad (45b)$$

$$X_0 \sim \mathcal{N}(m_0, \Sigma_0), \quad X_1 \sim \mathcal{N}(m_1, \Sigma_1). \qquad (45c)$$

As before, the above can be reformulated as an optimization (17) over $\mathcal{P}^u$ and can be solved using the proximal gradient algorithm with iteration (24). Again, each iteration of (24) is a covariance control problem for linear dynamics. However, the explicit realization of each iteration is slightly different.

The probability measure $\mathcal{P}_k^u$ is a Gaussian Markov process

$$dX_t = A_k(t)X_t dt + a_k(t)dt + \sqrt{\epsilon}g(t, z_t^k)dW_t. \qquad (46)$$

where the mean z^k and the covariance Σ^k satisfy

$$\dot{z}_t^k = A_k(t)z_t^k + a_k(t) \qquad (47a)$$

$$\dot{\Sigma}_t^k = A_k(t)\Sigma_t^k + \Sigma_t^k A_k(t)^T + \epsilon g(t, z_t^k)g(t, z_t^k)^T \quad (47b)$$

The above equations (46)-(47) should be compared with (37)-(38). The difference between them are due to the fact $g(t, x)$ depends on the value of state. In (46)-(47), we use the value of g at the mean z_t^k as an approximation.

Following the same analysis as in the Section IV, we obtain the explicit updates

$$A_{k+1}(t) = \frac{1}{1+\eta}[A_k(t) + \eta\hat{A}_k(t)] + g(t, z_t^k)K_k(t) \quad (48a)$$

$$a_{k+1}(t) = \frac{1}{1+\eta}[a_k(t) + \eta\hat{a}_k(t)] + g(t, z_t^k)d_k(t). \quad (48b)$$

Once $\{A_k, a_k\}$ converges to $(A^\star, a^\star)$, we can follow exactly the same steps as in Section IV-C (after replacing $B(t)$ by $g(t, z_t^\star)$) to retrieve the final optimal affine controller.

VI. NUMERICAL EXAMPLES

In this section we present several examples to illustrate our algorithm. All the experiments are carried out using a PC with 64GB RAM and an Intel Core i9 CPU.

A. Double integrator with drag

We consider the same nonlinear dynamical system used in [21]

$$dx_1 = x_2 dt, \qquad (49a)$$

$$dx_2 = (u - c_d\|x_2\|)dt + \sqrt{\epsilon}dW_t, \qquad (49b)$$

where c_d represents drag coefficient. To compare the proposed algorithm with the *iterative Covariance Steering (iCS)* algorithm [21], we conduct experiments for the two algorithms under the same settings but removing the chance constraints and trust region constraints in iCS. We choose $c_d = 0.005$ and $\epsilon = 0.1$, and also apply a time scaling to map time interval $[0, 1]$ to $[0, T]$. We choose the initial mean and covariance conditions as $m_0 =$

	iCS		ours	
Time discretization	25	50	25	**1000**
Solving time	126.1212	1528.1	**0.1184**	**0.2741**
Cost	31.3198	30.7929	2.9945	2.9283

TABLE I: Convergence time and cost comparison between the proposed method and the iCS algorithm proposed in [21] for the dynamics (49). The costs are computed using optimal controlled dynamics obtained from the two algorithms using Monte Carlo.

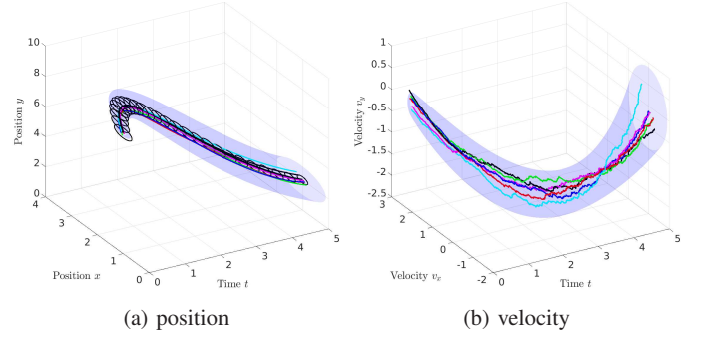

(a) position   (b) velocity

Fig. 1: Covariance and sampled position and velocity trajectories plot of the double integrator with drag dynamics (49), with a time scaling to $T = 5$ and time discretization $N = 1000$. We plot the covariances obtained from iCS for $N = 25$ in Fig. 1a in black ellipsoids. In contrast to iCS [21], our result satisfies the final time constraints.

$[1, 8, 2, 0]^T$, $\Sigma_0 = 0.01I$ and the target mean and covariance is chosen to be $m_T = [1, 2, -1, 0]^T$, $\Sigma_T = 0.1I$. In addition to control energy, in this experiment we add also a quadratic state cost $V(x_t) = (x_t - z_t^k)^T Q_t(x_t - z_t^k), Q_t = 0.1I$ in (9) as in [21]. The stopping criteria for the experiments is when the average over all time steps of relative Frobenius norm error of two consecutive A_k and a_k is less than a predefined threshold. Fig. 1 shows the resulting position and velocity distributions and sampled trajectories. We display the convergence time for different time discretizations in Tab. I. In the same settings, our proposed algorithm converges much faster (more than 1000 times speedup). We note that both algorithms converge to a locally minimum solution, and the reason for the convergence rate difference lies in that the baseline algorithm relies on optimization over discrete dynamics which leads to a constrained semi-definite program at each iteration while our proposed method optimizes directly over continuous dynamics at each iteration which has a closed-form solution. Our algorithm also has the advantage that its complexity increases linearly with time discretization. In terms of stability, our algorithm is guaranteed to have a solution that strictly satisfies the boundary conditions at each iteration if the linearized system is controllable.

B. 3-link manipulator

We next consider a 3-link manipulator example. By this example, we show that the proposed algorithm is effective in controlling a system governed by a manipulator equation in the form

$$M(q)\ddot{q} + C(q, \dot{q}) + g(q) = \tau + \sqrt{\epsilon}dW_t, \qquad (50)$$

where $q = [\theta_1, \theta_2, \theta_3]^T$ is the joint angles, $M(q)$ is the mass matrix, $C(q, \dot{q})$ is the Coriolis matrix, $g(q)$ is the gravity term and τ is the torque input applied on joints. Denote the state of the system as $X = [q, \dot{q}]^T$ and the input as $u = \tau$, then we have the nonlinear
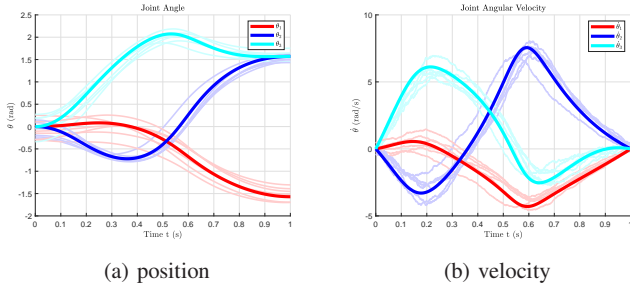
(a) position   (b) velocity

Fig. 2: Joint angle and velocity trajectory plot of the 3-DOF robot arm. The trajectories are simulated with full dynamics and noise injected from torque inputs. Mean trajectories of joint angles and velocities are highlighted.
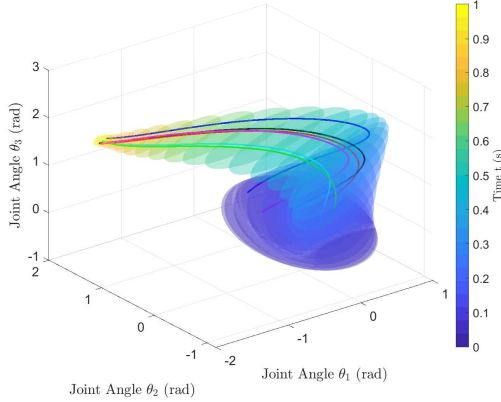

Fig. 3: 3σ covariance ellipsoid trajectory in joint space. Each joint angle ellipsoid is retrieved from covariance matrix Σ_t^k which satisfies (51b). The position of the ellipsoid is the mean z_t^k at time t.
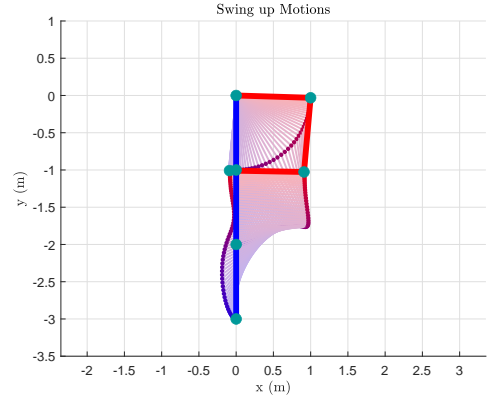

Fig. 4: Simulated trajectory of 3-DOF manipulator. The starting position is $X_0 = [0, 0, 0, 0, 0, 0]^T$ (hanging vertically, shown in blue), and the generated control policy steers the system to the vicinity of target state $X_T = [-\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, 0, 0, 0]^T$ (shown in red). The final position is not perfectly aligned with the target due to the stochasticity of the system.

control-affine system representation for the manipulator following (8)

$$dX_t = \begin{bmatrix} \dot{q} \\ M^{-1}(q)(-C(q,\dot{q}) - g(q)) \end{bmatrix} dt + \begin{bmatrix} 0 \\ M^{-1}(q)(u_t dt + \sqrt{\epsilon} dW_t) \end{bmatrix} \tag{51}$$

We choose the number of discretizations to 1000 with a time span $T = 1$ and noise intensity $\epsilon = 0.1$. We steer the system from initial mean $m_0 = [0, 0, 0, 0, 0, 0]^T$ and covariance $\Sigma_0 = 0.05I$, to $m_T = [-\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, 0, 0, 0]^T$ and covariance $\Sigma_T = 0.01I$. With convergence error set to 10^{-5}, it takes around 41 iterations and $6.2s$ on average to converge. Considering the underlying complexity of dynamics and the number of discretizations, the experiment shows the proposed algorithm is effective in solving the covariance steering problem for manipulator systems. The mean and sampled trajectories of the nonlinear stochastic system (51) under optimal feedback control policy are shown in Fig. 2. To illustrate the evolution of covariance and simulated trajectory, we show the covariance ellipsoid of joint angles in Fig. 3. The initial covariance shrinks as time goes on and reaches the target covariance. The simulated trajectories started in randomly sampled states from the initial distribution are bounded inside the 3σ-confident ellipsoid at each time stamp t. The trajectory simulation for 3-DOF manipulator is shown in Fig. 4.

VII. CONCLUSION

In this paper we presented a new approach to covariance steering problems for control-affine systems. We showed that this nonlinear covariance control problem can be reformulated as an optimization over the space of probability distributions of the trajectories. We then developed an efficient algorithm based on the proximal gradient algorithm. Each proximal iteration in the algorithm requires solving a linear covariance control problems, whose solutions exist in closed form. One distinguishing feature of this distributional formulation of nonlinear covariance steering and the resulting algorithm is that the optimal control is obtained for the continuous-time problems directly and the complexity of the algorithm is insensitive to the resolution of time-discretization. One potential future research direction is to extend the framework to account for hybrid dynamics.

APPENDIX

A. Proof of Lemma 2

By definition (19), the first order expansion of $F(\mathcal{P} + \delta\mathcal{P}) - F(\mathcal{P})$ is

$$\langle \frac{1}{\epsilon}\hat{V}(\mathcal{P} + \delta\mathcal{P}) - \log d\hat{\mathcal{P}}^0(\mathcal{P} + \delta\mathcal{P}), \mathcal{P} + \delta\mathcal{P}\rangle \tag{52}$$

$$-\langle \frac{1}{\epsilon}\hat{V}(\mathcal{P}) - \log d\hat{\mathcal{P}}^0(\mathcal{P}), \mathcal{P}\rangle$$

$$\approx \langle \frac{1}{\epsilon}\hat{V}(\mathcal{P}) - \log d\hat{\mathcal{P}}^0(\mathcal{P}), \delta\mathcal{P}\rangle$$

$$+\langle \frac{1}{\epsilon}\hat{V}(\mathcal{P} + \delta\mathcal{P}) - \frac{1}{\epsilon}\hat{V}(\mathcal{P}) - \log \frac{d\hat{\mathcal{P}}^0(\mathcal{P} + \delta\mathcal{P})}{d\hat{\mathcal{P}}^0(\mathcal{P})}, \mathcal{P}\rangle.$$

Here we use the notations $\hat{V}(\mathcal{P}), \hat{\mathcal{P}}^0(\mathcal{P})$ to emphasize the dependence of $\hat{V}, \hat{\mathcal{P}}^0$ on $\mathcal{P}$. By (15), $\langle \hat{V}(\mathcal{P} + \delta\mathcal{P}) - \hat{V}(\mathcal{P}), \mathcal{P}\rangle$ is approximately

$$\langle V(z_t + \delta z_t) - V(z_t) - \delta z_t^T \nabla V(z_t), \mathcal{P}\rangle$$

$$+ \langle \frac{1}{2}(x^T - z_t^T - \delta z_t^T)\nabla^2 V(z_t + \delta z_t)(x - z_t - \delta z_t)$$

$$- \frac{1}{2}(x^T - z_t^T)\nabla^2 V(z_t)(x - z_t), \mathcal{P}\rangle.$$

We only keep first order approximations and have used the fact that z_t is the mean of $\mathcal{P}$. The first term is clearly 0 after ignoring high

order information. The second term is approximately

$$\langle \frac{1}{2}(x^T - z_t^T)\nabla^2 V(z_t + \delta z_t)(x - z_t)$$
$$-\frac{1}{2}(x^T - z_t^T)\nabla^2 V(z_t)(x - z_t), \mathcal{P}\rangle$$
$$= \frac{1}{2}\int_0^1 [\text{Tr}(\nabla^2 V(z_t + \delta z_t)\Sigma_t) - \text{Tr}(\nabla^2 V(z_t)\Sigma_t)]dt$$
$$\approx \frac{1}{2}\int_0^1 [\delta z_t^T \nabla \text{Tr}(\nabla^2 V(z_t)\Sigma_t)]dt.$$

It follows that

$$\langle \hat{V}(\mathcal{P} + \delta\mathcal{P}) - \hat{V}(\mathcal{P}), \mathcal{P}\rangle \approx \frac{1}{2}\langle x^T \nabla \text{Tr}(\nabla^2 V(z_t)\Sigma_t), \delta\mathcal{P}\rangle. \quad (53)$$

By Girsanov theorem (10),

$$\frac{d\mathcal{P}}{d\hat{\mathcal{P}}^0}(\mathcal{P}) = \exp\left[\int_0^1 \frac{1}{2\epsilon}\|U\|^2_{(BB^T)^\dagger}dt + \frac{1}{\sqrt{\epsilon}}\langle U, BdW_t\rangle_{(BB^T)^\dagger}\right]$$

where

$$U = AX_t + a - \nabla f(z_t)^T X_t - f(z_t) + \nabla f(z_t)^T z_t.$$

A similar expression can be obtained for $d\mathcal{P}/d\hat{\mathcal{P}}^0(\mathcal{P} + \delta\mathcal{P})$. Combining them and using first order approximation yields

$$\langle \log \frac{d\hat{\mathcal{P}}^0(\mathcal{P})}{d\hat{\mathcal{P}}^0(\mathcal{P} + \delta\mathcal{P})}, \mathcal{P}\rangle \quad (54)$$
$$\approx \langle \frac{1}{2\epsilon}x^T \nabla \text{Tr}((BB^T)^\dagger(\nabla f(z_t)^T - A)\Sigma_t(\nabla f(z_t) - A^T)), \delta\mathcal{P}\rangle.$$

Plugging (53) and (54) into (52) yields (21).

REFERENCES

[1] A. Hotz and R. E. Skelton, "Covariance control theory," *International Journal of Control*, vol. 46, no. 1, pp. 13–32, 1987.

[2] T. Iwasaki and R. Skelton, "Quadratic optimization for fixed order linear controllers via covariance control," in *1992 American Control Conference*. IEEE, 1992, pp. 2866–2870.

[3] J.-H. Xu and R. Skelton, "An improved covariance assignment theory for discrete systems," *IEEE transactions on Automatic Control*, vol. 37, no. 10, pp. 1588–1591, 1992.

[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. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part I," *IEEE Trans. on Automatic Control*, vol. 61, no. 5, pp. 1158–1169, 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. on Automatic Control*, vol. 61, no. 5, pp. 1170–1180, 2016.

[7] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of inertial particles diffusing anisotropically with losses," in *Proc. American Control Conf.*, 2015, pp. 1252–1257.

[8] A. Halder and E. D. B. Wendel, "Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost," in *Proc. American Control Conf.*, 2016.

[9] Y. Chen, T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part III," *arXiv:1608.03622, IEEE Trans. on Automatic Control, to appear*, 2017.

[10] E. Bakolas, "Optimal covariance control for discrete-time stochastic linear systems subject to constraints," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 1153–1158.

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

[12] G. Zhu, K. M. Grigoriadis, and R. E. Skelton, "Covariance control design for hubble space telescope," *Journal of Guidance, Control, and Dynamics*, vol. 18, no. 2, pp. 230–236, 1995.

[13] J. Ridderhof and P. Tsiotras, "Uncertainty quantication and control during mars powered descent and landing using covariance steering," in *2018 AIAA Guidance, Navigation, and Control Conference*, 2018, p. 0611.

[14] K. Okamoto and P. Tsiotras, "Optimal stochastic vehicle path planning using covariance steering," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2276–2281, 2019.

[15] M. Kalandros, "Covariance control for multisensor systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1138–1157, 2002.

[16] M. Rudelson and R. Vershynin, "Hanson-Wright inequality and sub-Gaussian concentration," *Electronic Communications in Probability*, vol. 18, pp. 1–9, 2013.

[17] I. V. Girsanov, "On transforming a certain class of stochastic processes by absolutely continuous substitution of measures," *Theory of Probability & Its Applications*, vol. 5, no. 3, pp. 285–301, 1960.

[18] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[19] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.

[20] A. Beck, *First-order methods in optimization*. SIAM, 2017.

[21] J. Ridderhof, K. Okamoto, and P. Tsiotras, "Nonlinear uncertainty control with iterative covariance steering," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 3484–3490.

[22] Z. Yi, Z. Cao, E. Theodorou, and Y. Chen, "Nonlinear covariance control via differential dynamic programming," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3571–3576.

[23] E. Bakolas and A. Tsolovikos, "Greedy finite-horizon covariance steering for discrete-time stochastic nonlinear systems based on the unscented transform," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3595–3600.

[24] A. Tsolovikos and E. Bakolas, "Greedy covariance control for stochastic nonlinear systems using Gaussian process prediction models and the unscented transform," *arXiv preprint arXiv:2010.00778*, 2020.

[25] S. N. Cohen and R. J. Elliott, *Stochastic calculus and applications*. Springer, 2015, vol. 2.

[26] Q. Li, Y. Zhou, Y. Liang, and P. K. Varshney, "Convergence analysis of proximal gradient with momentum for nonconvex optimization," in *International Conference on Machine Learning*. PMLR, 2017, pp. 2111–2119.

[27] Y. Chen, T. T. Georgiou, and M. Pavon, "On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint," *Journal of Optimization Theory and Applications*, vol. 169, no. 2, pp. 671–691, 2016.

[28] S. Thijssen and H. Kappen, "Path integral control and state-dependent feedback," *Physical Review E*, vol. 91, no. 3, p. 032104, 2015.

[29] G. Williams, A. Aldrich, and E. A. Theodorou, "Model predictive path integral control: From theory to parallel computation," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 2, pp. 344–357, 2017.

[30] K. Caluya and A. Halder, "Wasserstein proximal algorithms for the schrödinger bridge problem: Density control with nonlinear drift," *IEEE Transactions on Automatic Control*, 2021.

[31] Y. Chen, T. T. Georgiou, and M. Pavon, "Fast cooling for a system of stochastic oscillators," *Journal of Mathematical Physics*, vol. 56, no. 11, p. 113302, 2015.

[32] K. F. Caluya and A. Halder, "Finite horizon density steering for multi-input state feedback linearizable systems," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3577–3582.

[33] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport in systems and control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, 2020.

[34] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*. Elsevier, 2014.

[35] V. Beneš, "Existence of optimal stochastic control laws," *SIAM Journal on Control*, vol. 9, no. 3, pp. 446–472, 1971.

[36] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.