

Optimal covariance control for partially observed linear stochastic systems under chance constraints

Dipankar Maity^{a,*}, Panagiotis Tsiotras^b

^a Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, Charlotte, 28223, NC, USA

^b Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, 30332, GA, USA

ARTICLE INFO

Keywords:

Chance constraints
Covariance steering
Innovation feedback
Optimal control
Stochastic control

ABSTRACT

We consider the optimal control problem of a partially observed discrete-time linear stochastic system with chance constraints on the state and control trajectories, while simultaneously controlling the first and second order moments of the state. Using a distribution-independent approximation of the chance constraints, we show that the class of affine controllers contains the optimal controller. Furthermore, this distribution-independent approximation is carried out while ensuring that any feasible solution for the approximated problem is also feasible for the original problem. Finally, we design a convex optimization problem to solve offline for the optimal controller for the approximated problem.

1. Introduction

While the majority of stochastic optimal control problems considers minimizing an expected state- and control-dependent cost, optimal covariance steering problem formulations [1,2] focus on steering the covariance of the final state to a prescribed value, while simultaneously minimizing an expected state- and control-dependent cost. As a result, the controller solving the covariance steering problem not only optimally steers the state with minimum (state and control related) cost, but it also minimizes the *uncertainty* in the terminal state, measured in terms of its mean and covariance. In this spirit, the optimal covariance steering problem is closely related to the recent developments in distributional control [3–6] where the objective is to steer the distribution of the state of a dynamical system from a given initial distribution to a given final distribution, establishing a connection between optimal control and Schrödinger bridges [7,8].

To the best of our knowledge, the problem of controlling the state covariance was first introduced in [1,2], and it was later extended in several works including [9–12]. In all these prior studies, the emphasis is placed on infinite horizon problems. The objective of these works is to steer the state covariance of a stochastic linear system to a prescribed steady-state value and stay there. Recently, the focus has shifted to the finite horizon case [13–15], where the covariance of a linear system is controlled to be within a given set while minimizing an expected quadratic cost.

Recent studies, such as [16], have also looked into optimal covariance steering control problems with chance constraints by imposing

restrictions on the maximum probability of chance constraint violation. Such chance constraints are generally defined for the entire time horizon rather than only at a terminal time, and thus, they are different from previous studies in the covariance control community [3, 6,14]. In [17,18], chance-constrained optimal control problems were analyzed to guarantee performance under uncertainty. Furthermore, this framework has been extended to stochastic model predictive control [19,20], vehicle path planning [21,22], and spacecraft control [23] problems. Extensions include incorporating (decoupled) state and input constraints [24], and to non-linear systems [25] using successive approximations. In [26] a model predictive path integral approach in conjunction to covariance steering has been used for trajectory optimization of stochastic linear systems.

The majority of all these prior approaches assume perfect sensing and full state observation. An exception is [27], which considers a partial and noisy observation model. However, [27] imposes a linear structure on the controller and does not incorporate any chance constraints. Instead, it introduces a quadratic (convex) constraint that is structurally similar to the quadratic cost function. The combination of convex cost and constraint functions, along with the restriction to linear controllers, results in a convex optimization problem that can be solved efficiently. In this article, we consider a chance-constrained covariance steering optimal control problem for discrete-time linear systems with partial noisy observations. A similar problem from the stochastic model predictive community has been studied in [28], which, however, does not provide any optimality guarantees on the controller structure. Our problem is similar in spirit with [29], where an approximated

* Corresponding author.

E-mail addresses: dmaity@charlotte.edu (D. Maity), tsiotras@gatech.edu (P. Tsiotras).

version of the chance constraint problem is studied under the restrictive assumption that the admissible controller is linear. It is, however, different from [29] in terms of the generality of the imposed chance constraints; more importantly, in this paper we do not adhere to any specific structure of the optimal controller.

The contributions of this article are as follows: First, we propose a *sufficient approximation* of the output feedback covariance steering problem under chance constraints, that is, any feasible solution of the proposed relaxed problem is also feasible for the original problem. The (sufficient) approximation is distribution-independent, that is, the approximation is valid regardless of the structure of the control law. Next, for the approximated problem, we prove that the optimal controller lies in the space of affine controllers. Finally, we construct a convex optimization problem to efficiently solve for the optimal covariance steering controller.

2. Problem formulation

Let us consider the linear, discrete-time stochastic dynamical system,

$$X_{t+1} = A_t X_t + B_t U_t + W_t, \quad (1a)$$

$$Y_t = C_t X_t + V_t, \quad (1b)$$

$$Z_t = F_t X_t + G_t U_t, \quad (1c)$$

where $X_t \in \mathbb{R}^n$, $Y_t \in \mathbb{R}^p$, $Z_t \in \mathbb{R}^\ell$, and $U_t \in \mathbb{R}^m$ are the state, the sensor measurement, the output, and the control input signals, respectively. The process and measurement noises are $\{W_t\}_{t \in \mathbb{N}_0}$ and $\{V_t\}_{t \in \mathbb{N}_0}$ ($\mathbb{N}_0$ is the set of non-negative integers), respectively, where¹, for all $t, s, k \in \mathbb{N}_0$, with $t \neq s$, $W_t \perp W_s$, $W_t \perp V_k$, $V_t \perp V_s$, $W_t \perp X_0$, and $V_t \perp X_0$. The primitive random variables W_t , V_t and X_0 are distributed as $W_t \sim \mathcal{N}(0, W_t)$, $V_t \sim \mathcal{N}(0, V_t)$ and $X_0 \sim \mathcal{N}(\mu_0, X_0)$, respectively. For notational compactness, we will also denote $X_0 = \mu_0 + W_{-1}$ where $W_{-1} \sim \mathcal{N}(0, X_0)$.

The finite horizon chance-constrained optimal control problem is given below.

Problem 1. Given the dynamics, measurement, and output Eqs. (1), find a control law to solve the following optimization problem

$$\min_U J(U) \triangleq \mathbb{E} \left[\sum_{t=0}^T Z_t^\top Q_t Z_t \right], \quad (2a)$$

$$\text{subject to} \quad \mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1] \leq \delta_t, \quad t = 0, \dots, T-1, \quad (2b)$$

$$\mathbb{E}[Z_T] = \mu_f, \quad \mathbb{E}[(Z_T - \mu_f)(Z_T - \mu_f)^\top] \leq \Sigma_f, \quad (2c)$$

where $U = \{U_0, U_1, \dots, U_{T-1}\}$ is the entire control sequence over the horizon of length T , and for any vector v and any compatible matrix M , we use $\|v\|_M^2$ to denote the scalar $v^\top M v$.

Notice that for a special choice of the parameters F_t , G_t and Q_t , the cost function can be expressed as $\mathbb{E}[\sum_{t=0}^T X_t^\top L_t X_t + U_t^\top R_t U_t]$, which has been considered in prior works [16,29]. While prior works consider the constraints and the cost to be decoupled in the states and controls, our formulation encapsulates coupling between the states and inputs using the output variable Z_t .

The parameters Q_t , r_t , S_t , δ_t , μ_f and Σ_f are user-defined and fixed *a priori* for the optimization Problem 1. The probabilistic constraints in (2c) encode the terminal constraints on the mean and the covariance of the output Z_t , whereas, the constraints in (2b) ensure that Z_t at time t remains within a given (safe) ellipsoid² with a probability no less than $1 - \delta_t$. The chance constraints in (2b) can also be viewed

as an intent to confine the output trajectory $\{Z_0, \dots, Z_{T-1}\}$ within a neighborhood (determined by the matrix S_t) of a predefined reference trajectory $\{r_0, \dots, r_{T-1}\}$ with a prescribed probability (controlled by δ_t). In the problem formulation, the safe regions (determined by S_t and r_t) along with the constraint violation probabilities (δ_t) are time-varying.

The feasibility of Problem 1 depends on the parameters (e.g., Q_t , r_t , S_t , δ_t , μ_f , etc.) of the problem. For certain choices of these parameters, Problem 1 may become infeasible. In the subsequent analysis, we assume that Problem 1 is feasible, and furthermore, we also assume that $G_t^\top Q_t G_t > 0$ for all t , which is a standard assumption³ in LQG optimal control.

The chance-constrained optimization in Problem 1 is analytically intractable, in general, due to the constraint $\mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1] < \delta_t$. In this article, we propose a relaxation to this constraint and solve the relaxed problem instead. Prior works [22,24,28,29] on chance-constrained optimal covariance steering have also proposed approximations to the chance constraints to circumvent the intractability imposed by this constraint. However, the approximations proposed in those works are valid only if the state X_t has a certain density function or the controller is constrained to be of a very specific structure [22,23,29]. The approximations and the overall framework proposed in these prior works fail if the problem is studied under a different class of controllers, or when the actual distribution of the state is no longer the assumed one. In this article, we propose a relaxation that is valid for *any* admissible control law, thus, enabling us to consider a much larger class of controllers for solving the relaxed optimization problem.

2.1. A relaxed formulation

In this section, we develop a relaxation to Problem 1 by replacing the chance constraint $\mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1] \leq \delta_t$ with a new set of constraints that are analytically tractable. First, notice that the satisfaction (or violation) of the constraint (2b) depends on the probability distribution of the random variable Z_t . At this point, we cannot make any assumptions on the probability distribution of Z_t since it depends on the control law yet to be designed. Therefore, the relaxation of the constraint (2b) has to be valid for any distribution of Z_t , i.e., we need a *distribution independent* relaxation of the chance constraint (2b). In this article, we use Markov's inequality⁴ to obtain a distribution-independent approximation of the quantity $\mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1]$. In particular, Markov's inequality for the random variable $\|Z_t - r_t\|_{S_t}^2$ yields,

$$\mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1] \leq \mathbb{E}[\|Z_t - r_t\|_{S_t}^2]. \quad (3)$$

We therefore relax the chance constraint (2b) as

$$\mathbb{E}[\|Z_t - r_t\|_{S_t}^2] \leq \delta_t. \quad (4)$$

Notice that the new constraint $\mathbb{E}[\|Z_t - r_t\|_{S_t}^2] \leq \delta_t$ is a *sufficient approximation* of the chance constraint (2b), since any system trajectory satisfying the constraint (4) also satisfies the original constraint (2b). The relaxed optimization problem we will consider for the subsequent analysis is presented in Problem 2.

Problem 2. A relaxed version of Problem 1 is given as follows

$$\min_U \bar{J}(U) \triangleq \mathbb{E} \left[\sum_{t=0}^T Z_t^\top Q_t Z_t \right], \quad (5a)$$

$$\text{subject to} \quad \mathbb{E}[\|Z_t - r_t\|_{S_t}^2] \leq \delta_t, \quad t = 0, \dots, T-1, \quad (5b)$$

$$\mathbb{E}[Z_T] = \mu_f, \quad \mathbb{E}[(Z_T - \mu_f)(Z_T - \mu_f)^\top] \leq \Sigma_f. \quad (5c)$$

¹ $X \perp Y$ denotes that the random variables X and Y are independent.

² In this article, an ellipsoid $\mathcal{E}(q, Q)$ in $\mathbb{R}^n$ with center q and shape matrix Q denotes the set $\{x \in \mathbb{R}^n \mid (x - q)^\top Q (x - q) \leq 1\}$.

³ This is equivalent to the assumption that the control weight matrix R_t is positive definite in the cost function $\mathbb{E}[\sum_{t=0}^T X_t^\top L_t X_t + U_t^\top R_t U_t]$

⁴ Markov's inequality states that given a non-negative random variable X and a real positive scalar $a > 0$, then $\mathbb{P}(X \geq a) \leq \mathbb{E}[X]/a$.

In what follows, we will focus on finding an optimal controller to solve [Problem 2](#). It is important to mention here that other approximations to (2b) may be tighter, however, such approximations may lead to analytically intractable problems and may also depend on the particular distribution of the state. They are thus beyond the scope of this paper.

Before proceeding, let us discuss an innovation process which will be instrumental in solving [Problem 2](#).

3. The innovation process

At time t , the controller receives the measurements $Y_{0:t} \triangleq \{Y_0, \dots, Y_t\}$, based on which a control action U_t is computed. Thus, the control input U_t at time t is adapted to the σ -field $\sigma(Y_{0:t})$. One can verify that the observations $Y_{0:T}$ are correlated and depend on the control inputs $U_{0:T-1}$. In what follows, we will perform a whitening transformation to the observations in order to progressively construct a sequence of signals $\xi_{0:T-1}$, called the “innovation sequence”, such that $\xi_t \perp \xi_s$ for all $t \neq s$. Furthermore, in the process of generating $\xi_{0:T-1}$ from $Y_{0:T-1}$, no information is lost [30]. We formally present this result in [Lemma 1](#).

Lemma 1. *The σ -algebra $\sigma(Y_{0:t})$ generated by the observation process is the same as the σ -algebra $\sigma(\xi_{0:t})$ generated by the innovation process $\{\xi_0, \dots, \xi_t\}$, where*

$$\begin{aligned} \xi_t &= Y_t - \mathbb{E}[Y_t | Y_{0:t-1}, U_{0:t-1}], \\ \xi_0 &\triangleq Y_0 - \mathbb{E}[Y_0] = C_0(X_0 - \mu_0) + V_0, \end{aligned} \quad (6)$$

where $Y_{0:t-1}$ and $U_{0:t-1}$ are the observation and control histories, respectively.

Proof. The proof of Lemma 6 is presented in [Appendix A.1](#).

The purpose of introducing the innovation process is to present a tractable computational approach to solve [Problem 2](#), as well as to greatly simplify the analysis in the subsequent sections. Specifically, we will make use of the fact that the distribution of the innovation signal is control independent and can be computed *a priori* using the next lemma, whereas, the measurement process Y_t does not have such a property.

Lemma 2. *For all t , ξ_t is a Gaussian random variable with zero mean and covariance M_t such that*

$$\begin{aligned} M_{t+1} &= C_{t+1} \Sigma_{t+1|t} C_{t+1}^\top + \mathcal{V}_{t+1}, \\ \Sigma_{t+1|t} &= A_t \Sigma_t A_t^\top + \mathcal{W}_t, \\ \Sigma_{0|-1} &= \mathcal{X}_0, \\ \Sigma_{t+1} &= \Sigma_{t+1|t} - \Sigma_{t+1|t} C_{t+1}^\top M_{t+1}^{-1} C_{t+1} \Sigma_{t+1|t}, \end{aligned}$$

Moreover, the sequence of random variables $\{\xi_0, \dots, \xi_t\}$ is uncorrelated for all t .

Proof. The proof of this lemma can be constructed in a similar way as the one presented in [30]. However, for the sake of completeness, we present it in [Appendix A.2](#).

The matrix $\mathcal{X}_0$ in [Lemma 2](#) is the covariance of the initial state X_0 , and $\mathcal{W}_t, \mathcal{V}_t$ are the covariances of the process noise W_t and the measurement noise V_t , respectively. [Lemma 2](#) is equivalent to the following facts: (a) the innovation sequence $\xi_{0:t}$ does not depend on the control history $U_{0:t}$ for all $t = 0, \dots, T$, and (b) the innovation sequence is a Gaussian uncorrelated noise sequence with zero mean and covariance M_t . Consequently, $\xi_t \perp \xi_s$ for all $s \neq t$. It also follows from [Lemma 1](#) that the $\sigma(Y_{0:t})$ -measurable class of controls is the same class of controls that are $\sigma(\xi_{0:t})$ -measurable. Thus, without any loss of generality, in the subsequent analysis we will restrict ourselves to the class of $\sigma(\xi_{0:t})$ -measurable functions to find an optimal controller. We conclude this section with the following remark on the properties of the innovation signal for a perfectly observed system.

Remark 1. Under the perfect observation case, that is, $C_t = I$ and $\mathcal{V}_t = 0$, we have that $\xi_t = W_{t-1}$, $\Sigma_t = 0$, and $M_t = \Sigma_{|t-1} = \mathcal{W}_{t-1}$, where $\mathcal{W}_{t-1}$ is the covariance of the process noise W_{t-1} .

4. Equivalent fully observed problem

In this section, we show that the partially observed system (1) can be reduced to a *fully observed* system, which will further reduce the complexity of the subsequent analysis. To this end, we introduce two filtered signals and use them to recast the partially observed system (1) as an equivalent perfectly observed system. Let us define the following quantities

$$\hat{X}_t \triangleq \mathbb{E}[X_t | \xi_{0:t}, U_{0:t-1}], \quad (7a)$$

$$\hat{Z}_t \triangleq \mathbb{E}[Z_t | \xi_{0:t}, U_{0:t-1}] = F_t \hat{X}_t + G_t U_t, \quad (7b)$$

$$\tilde{X}_t \triangleq X_t - \hat{X}_t, \quad (7c)$$

$$\tilde{Z}_t \triangleq Z_t - \hat{Z}_t, \quad (7d)$$

$$\hat{X}_0 = \Sigma_0 C_0^\top (C_0 \Sigma_0 C_0^\top + \mathcal{V}_0)^{-1} \xi_0, \quad (7e)$$

Using the definition of $\hat{X}_t$ from (7), it can be verified that

$$\hat{X}_{t+1} = A_t \hat{X}_t + B_t U_t + K_{t+1} \xi_{t+1}, \quad (8)$$

where $K_t = \Sigma_{|t-1} C_t^\top M_t^{-1}$ is the *Kalman gain*.

We may further write

$$\mathbb{E}[Z_t^\top Q_t Z_t] = \mathbb{E}[\hat{Z}_t^\top Q_t \hat{Z}_t] + \mathbb{E}[\tilde{Z}_t^\top Q_t \tilde{Z}_t],$$

and

$$\mathbb{E}[\|Z_t - r_t\|_{S_t}^2] = \mathbb{E}[\|\hat{Z}_t - r_t\|_{S_t}^2] + \mathbb{E}[\|\tilde{Z}_t\|_{S_t}^2].$$

Notice that $\tilde{X}_t$ and $\tilde{Z}_t$ do not depend on the control signal since their dynamics can be written as

$$\begin{aligned} \tilde{X}_{t+1} &= A_t \tilde{X}_t + W_t - K_{t+1} \xi_{t+1}, \\ \tilde{Z}_t &= F_t \tilde{X}_t, \quad \tilde{X}_{-1} = 0, \end{aligned}$$

where we introduce the initial condition $\tilde{X}_{-1} = 0$ so that $\tilde{X}_0 = A_{-1} \tilde{X}_{-1} + W_{-1} - K_0 \xi_0 = X_0 - K_0 \xi_0$ with $A_{-1} = I$. One can also verify that $\mathbb{E}[\tilde{X}_t \tilde{X}_t^\top] = \Sigma_t$ where Σ_t is defined in [Lemma 2](#), and likewise, $\mathbb{E}[\tilde{Z}_t \tilde{Z}_t^\top] = F_t \Sigma_t F_t^\top$. To this end, we rewrite [Problem 2](#) as follows.

Problem 3. An equivalent optimization problem to [Problem 2](#) is

$$\min_{U^\star} \bar{J}(U) = \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^\top Q_t \hat{Z}_t \right] + c_1, \quad (9)$$

$$\text{subject to} \quad \mathbb{E}[\|\hat{Z}_t - r_t\|_{S_t}^2] \leq \tilde{\delta}_t, \quad t = 0, \dots, T-1, \quad (10)$$

$$\mathbb{E}[\hat{Z}_T] = \mu_f, \quad \mathbb{E}[(\hat{Z}_T - \mu_f)(\hat{Z}_T - \mu_f)^\top] \leq \tilde{\Sigma}_f. \quad (11)$$

where $c_1 = \sum_{t=0}^T \text{tr}(Q_t F_t \Sigma_t F_t^\top)$, $\tilde{\delta}_t = \delta_t - \text{tr}(F_t^\top S_t F_t \Sigma_t)$ and $\tilde{\Sigma}_f = \Sigma_f - F_T \Sigma_T F_T^\top$.

The quantity c_1 in [Problem 3](#) does not depend on the control law and thus it can be omitted from (9) without affecting the optimal solution.

While X_t (and hence Z_t) is partially observed through Y_t , the variable $\hat{X}_t$ (and hence $\hat{Z}_t$) is fully observed, i.e., the controller can compute $\hat{X}_t$ exactly, whereas, it cannot compute X_t exactly. The emergence of the fully observed system $(\hat{X}_t, Y_t, \hat{Z}_t)$ from the partially observed system (X_t, Y_t, Z_t) is due to the following corollary.

Corollary 1. *Using $\hat{X}_t$, the partially observed system (1) for [Problem 2](#) becomes a perfectly observed system for [Problem 3](#) as described below*

$$\hat{X}_{t+1} = A_t \hat{X}_t + B_t U_t + K_{t+1} \xi_{t+1}, \quad (12a)$$

$$O_t = \hat{X}_t, \quad (12b)$$

$$\hat{Z}_t = F_t \hat{X}_t + G_t U_t. \quad (12c)$$

The controller cannot compute X_t since it does not have access to the random variables $X_0, W_0, \dots, W_{t-1}$. However, using the innovation signals ξ_t , the controller can compute $\hat{X}_t$ by following the dynamics (12a). In (12a), $K_{t+1}\xi_{t+1}$ emulates the noise W_t that affects the system dynamics (1a). As mentioned in Remark 1, for the perfect observation case ($C_t = I, \mathcal{V}_t = 0$ for all t) we obtain $\xi_{t+1} = W_t$, and $M_{t+1} = \Sigma_{t+1|t}$, which further implies that $K_{t+1} = \Sigma_{t+1|t} C_{t+1}^\top M_{t+1}^{-1} = \Sigma_{t+1|t} \Sigma_{t+1|t}^{-1} = I$. Consequently, for the perfect observation case we obtain that $K_{t+1}\xi_{t+1} = W_t$ and $\hat{X}_t = X_t$, for all t .

Notice from the expressions of $c_1, \delta_t, \bar{\Sigma}_f$ that partial observation makes the problem more constrained and increases the optimal cost. The three quantities c_1, δ_t and $\bar{\Sigma}_f$ are directly affected by the parameters of the measurement model, namely, C_t and $\mathcal{V}_t$. Since $\delta_t \geq 0$ and $\bar{\Sigma}_f \geq 0$ are necessary for the feasibility of Problem 3, we can derive sufficient conditions for Problem 3 to be infeasible by finding conditions on these parameters such that either $\bar{\Sigma}_f < 0$ or $\delta_t < 0$ for some t . These sufficient conditions for the infeasibility of Problem 3 are provided below.

Fact 1. *Either of the following two conditions renders Problem 3 infeasible:*

$$(i) F_T \Sigma_T F_T^\top > \Sigma_f, \quad (13a)$$

$$(ii) \text{ There exist } t \in \{0, \dots, T-1\} \text{ such that } \text{tr}(F_t \Sigma_t F_t^\top S_t) > \delta_t. \quad (13b)$$

The sufficient conditions (13) for the infeasibility of Problem 3 are easy to check and do not require the solution of an optimization problem. It is worth noticing that these conditions depend on all the problem parameters except $r_{0:T-1}$ and μ_f . Between the two sufficient conditions of infeasibility in (13), the condition $\Sigma_f < F_T \Sigma_T F_T^\top$ is fundamental to the original and the relaxed problems, that is, it ensures infeasibility for any of the Problems 1–3. On the other hand, the condition $\delta_t = \delta_t - \text{tr}(F_t^\top S_t F_t S_t) < 0$ for some t ensures infeasibility of Problems 2–3, whereas, the original problem (Problem 1) can still be feasible.

In the next section, we design the optimal controller for Problem 3 under the assumption that the problem parameters are chosen such that none of the conditions in (13) is satisfied, i.e., Problem 3 is feasible.

5. Optimal controller synthesis

In this section, we construct the optimal controller corresponding to Problem 3. The main result in this section is Theorem 1, where we show that the optimal control law has an affine structure. Specifically, we show that for every nonlinear controller, there is an affine controller that performs equally well with respect to the objective (9), and satisfies the constraints (10)–(11). Before stating Theorem 1, we provide an important lemma that is instrumental for proving Theorem 1.

Lemma 3. *Let γ be any feasible (randomized) linear control law for Problem 3 of the form*

$$U_t = \gamma_t(\hat{X}_{0:t}, U_{0:t-1}) = \bar{\gamma}_t(\hat{X}_{0:t}, U_{0:t-1}) + \zeta_t, \quad (14)$$

where $\bar{\gamma}$ is a deterministic linear function and ζ_t is a random variable such that $\zeta_t \perp \zeta_s$ for all $t \neq s$ and $\zeta_t \perp \xi_k$ for all t and k . Then, the deterministic (linear) control law γ^d such that

$$U_t = \gamma_t^d(\hat{X}_{0:t}, U_{0:t-1}) \triangleq \bar{\gamma}_t(\hat{X}_{0:t}, U_{0:t-1}) + \mathbb{E}[\zeta_t], \quad (15)$$

is also feasible for Problem 3 and yields a cost no worse than the one obtained by γ .

Proof. Due to the linearity of $\bar{\gamma}$, we obtain $\mathbb{E}[\bar{\gamma}_t(\hat{X}_{0:t}, U_{0:t-1})] = \bar{\gamma}_t(\mathbb{E}[\hat{X}_{0:t}], \mathbb{E}[U_{0:t-1}])$. Now, let $\hat{X}_t^\gamma$ denote the state evolution under the randomized law γ and let $\hat{X}_t^d$ denote the same under the deterministic law γ^d . One can verify that $\mathbb{E}[\hat{Z}_t^\gamma] = \mathbb{E}[\hat{Z}_t^d]$ and $\text{Var}(\hat{Z}_t^\gamma) \geq \text{Var}(\hat{Z}_t^d)$ for

all t , where $\text{Var}(\cdot)$ denotes the variance⁵ of a random variable. Based on these two observations, it follows that, if γ satisfies the constraints (10)–(11), then γ^d does so too. Therefore, γ^d is a feasible solution to Problem 3.

Under the control law γ^d , we may rewrite the objective in (9) as follows,

$$\begin{aligned} \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^{d\top} Q_t \hat{Z}_t^d \right] &= \sum_{t=0}^T [\mathbb{E}[\hat{Z}_t^d]^\top Q_t \mathbb{E}[\hat{Z}_t^d] + \text{tr}(\text{Var}(\hat{Z}_t^d) Q_t)] \\ &\stackrel{(a)}{\leq} \sum_{t=0}^T [\mathbb{E}[\hat{Z}_t^\gamma]^\top Q_t \mathbb{E}[\hat{Z}_t^\gamma] + \text{tr}(\text{Var}(\hat{Z}_t^\gamma) Q_t)] \\ &= \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^{\gamma\top} Q_t \hat{Z}_t^\gamma \right], \end{aligned}$$

where in deriving (a), we have used the facts that $\mathbb{E}[\hat{Z}_t^\gamma] = \mathbb{E}[\hat{Z}_t^d]$ and $\text{Var}(\hat{Z}_t^\gamma) \geq \text{Var}(\hat{Z}_t^d)$ for all t . Therefore, γ^d results in a cost no higher than the one produced by γ . This completes the proof.

Lemma 3 implies that the class of randomized controllers of the form (14) does not produce a better cost than the class of deterministic controllers of the form (15). Such an observation is ubiquitous in the context of a Markov-decision-process (MDP), where one can show that for any randomized policy there exists a deterministic policy that performs equally well. We are now ready to state the main result of this paper.

Theorem 1. *The optimal solution of Problem 3 is of the form*

$$U_t = L_t \hat{X}_t + \bar{u}_t, \quad (16)$$

where $L_t \in \mathbb{R}^{m \times n}$ and $\bar{u}_t \in \mathbb{R}^m$.

Proof. The proof is based on the fact that for every nonlinear controller, we can construct a feasible affine controller that produces a cost that is less than or equal to the one obtained using the nonlinear controller.

For any control policy $\gamma = \{\gamma_0, \gamma_1, \dots, \gamma_T\}$, we explicitly write $\hat{X}_t^\gamma$ and U_t^γ to denote the state and control input at time t obtained by using the policy γ . We now show that for any non-linear control policy $\gamma_t : \{\hat{X}_{0:t}^\gamma, U_{0:t-1}^\gamma\} \rightarrow \mathbb{R}^m$, there exists an affine policy $\ell = \{\ell_0, \ell_1, \dots, \ell_T\}$ such that $U_t^\ell = \ell_t(\hat{X}_{0:t}^\gamma, U_{0:t-1}^\gamma) \triangleq L_t \hat{X}_t^\gamma + \bar{u}_t^\ell$ performs no worse than the non-linear policy γ for Problem 3.

To this end, let $\mathbb{P}_\gamma(\hat{X}_{0:T}^\gamma, U_{0:T-1}^\gamma)$ be the state-control distribution generated by the control policy γ . Let us consider a candidate affine control ℓ_t of the form

$$\begin{aligned} U_t^\ell &= \ell_t(\hat{X}_t^\gamma) \\ &= \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma) \text{Var}_\gamma(\hat{X}_t^\gamma)^{-1} (\hat{X}_t^\gamma - \mathbb{E}_\gamma[\hat{X}_t^\gamma]) + \zeta_t, \end{aligned} \quad (17)$$

where $\zeta_t \sim \mathcal{N}(\bar{\zeta}_t, \Lambda_t)$ is a Gaussian random variable with mean $\bar{\zeta}_t$ and variance Λ_t . The expressions of $\bar{\zeta}_t$ and Λ_t are as follows,

$$\bar{\zeta}_t = \mathbb{E}_\gamma[U_t^\gamma], \quad (18)$$

$$\Lambda_t = \text{Var}_\gamma(U_t^\gamma) - \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma) \text{Var}_\gamma(\hat{X}_t^\gamma)^{-1} \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma)^\top, \quad (19)$$

where $\mathbb{E}_\gamma[\cdot]$, $\text{Var}_\gamma(\cdot)$ and $\text{Cov}_\gamma(\cdot)$ denote the expectation, variance and covariance, respectively, under the distribution $\mathbb{P}_\gamma$. The random variable ζ_t is chosen such that it satisfies $\zeta_t \perp \zeta_s$ for all $s \neq t$, and $\zeta_t \perp \xi_k$ for all t, k .

Let $\mathbb{P}_\ell(\hat{X}_{0:T}^\ell, U_{0:T-1}^\ell)$, $\mathbb{E}_\ell[\cdot]$, Var_ℓ , and Cov_ℓ denote the distribution, expectation, variance and covariance, respectively, induced by the affine control law (17). We will show that if γ is a feasible control law, then ℓ defined by (17) is also feasible, and produces a cost no worse

⁵ For any two random variables X and Y , we use the notation $\text{Var}(X) \triangleq \mathbb{E}[X X^\top] - \mathbb{E}[X] \mathbb{E}[X]^\top$ and $\text{Cov}(X, Y) \triangleq \mathbb{E}[X Y^\top] - \mathbb{E}[X] \mathbb{E}[Y]^\top = \text{Cov}(Y, X)^\top$

than the one produced by γ . We prove this statement in three steps, where the goal is to show that the first and second order moments of the state and the control are the same for both the nonlinear and the affine controller.

Step 1. We show using an inductive argument that, under the control law ℓ , we have $\mathbb{E}_\gamma[\hat{X}_t^\gamma] = \mathbb{E}_\ell[\hat{X}_t^\ell]$ and $\mathbb{E}_\gamma[U_t^\gamma] = \mathbb{E}_\ell[U_t^\ell]$ for all t .

The policy γ or ℓ does not affect $\hat{X}_0$, and hence, $\hat{X}_0^\gamma = \hat{X}_0^\ell$,⁶ and also $\mathbb{P}_\ell(\hat{X}_0) = \mathbb{P}_\gamma(\hat{X}_0) \sim \mathcal{N}(\mu_0, K_0 M_0 K_0^\top)$. Therefore, $\mathbb{E}_\ell[\hat{X}_0^\ell] = \mu_0 = \mathbb{E}_\gamma[\hat{X}_0^\gamma]$. Furthermore, from (17), $\mathbb{E}_\ell[U_0^\ell] = \mathbb{E}_\ell[\ell_0(\hat{X}_0^\ell)] = \mathbb{E}_\ell[\zeta_0] = \mathbb{E}_\gamma[U_0^\gamma]$ since $\zeta_0 \sim \mathcal{N}(\bar{\zeta}_0, A_0)$ and $\bar{\zeta}_0 = \mathbb{E}_\gamma[U_0^\gamma]$.

Let us now assume that, for some time k , we have $\mathbb{E}_\ell[\hat{X}_k^\ell] = \mathbb{E}_\gamma[\hat{X}_k^\gamma]$ and $\mathbb{E}_\ell[U_k^\ell] = \mathbb{E}_\gamma[U_k^\gamma]$. Therefore, at time $k+1$,

$$\begin{aligned} \mathbb{E}_\ell[\hat{X}_{k+1}^\ell] &= \mathbb{E}_\ell[A_k \hat{X}_k^\ell + B_k U_k^\ell + K_{k+1} \xi_{k+1}] \\ &= A_k \mathbb{E}_\ell[\hat{X}_k^\ell] + B_k \mathbb{E}_\ell[U_k^\ell] \\ &= A_k \mathbb{E}_\gamma[\hat{X}_k^\gamma] + B_k \mathbb{E}_\gamma[U_k^\gamma] = \mathbb{E}_\gamma[\hat{X}_{k+1}^\gamma], \end{aligned}$$

and from (17), $\mathbb{E}_\ell[U_{k+1}^\ell] = \mathbb{E}_\ell[\zeta_{k+1}] = \mathbb{E}_\gamma[U_{k+1}^\gamma]$. Thus, for all t , $\mathbb{E}_\gamma[\hat{X}_t^\gamma] = \mathbb{E}_\ell[\hat{X}_t^\ell]$ and $\mathbb{E}_\gamma[U_t^\gamma] = \mathbb{E}_\ell[U_t^\ell]$.

Step 2. Next, we show that $\text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_t^\gamma \\ U_t^\gamma \end{bmatrix} \right) = \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_t^\ell \\ U_t^\ell \end{bmatrix} \right)$ for all t . Since the control law ℓ is of the form (17), we may write $U_t^\ell = \bar{u}_t + L_t \hat{X}_t^\ell + \zeta_t$ where $\bar{u}_t = -\text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma) \text{Var}_\gamma(\hat{X}_t^\gamma)^{-1} \mathbb{E}_\gamma[\hat{X}_t^\gamma]$ and $L_t = \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma) \text{Var}_\gamma(\hat{X}_t^\gamma)^{-1}$ are deterministic quantities. Therefore, we may further write

$$\begin{aligned} \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_t^\ell \\ U_t^\ell \end{bmatrix} \right) &= \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_t^\ell \\ L_t \hat{X}_t^\ell + \zeta_t \end{bmatrix} \right) \\ &= \begin{bmatrix} \text{Var}_\ell(\hat{X}_t^\ell) & \text{Var}_\ell(\hat{X}_t^\ell) L_t^\top \\ L_t \text{Var}_\ell(\hat{X}_t^\ell) & L_t \text{Var}_\ell(\hat{X}_t^\ell) L_t^\top + \text{Var}_\ell(\zeta_t) \end{bmatrix}, \end{aligned}$$

since ζ_t is a random variable independent of $\hat{X}_t^\ell$.

At time $t = 0$, we have $\text{Var}_\ell(\hat{X}_0^\ell) = \text{Var}_\gamma(\hat{X}_0^\gamma) = K_0 M_0 K_0^\top$, and therefore

$$\begin{aligned} \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_0^\ell \\ U_0^\ell \end{bmatrix} \right) &= \begin{bmatrix} \text{Var}_\ell(\hat{X}_0^\ell) & \text{Var}_\ell(\hat{X}_0^\ell) L_0^\top \\ L_0 \text{Var}_\ell(\hat{X}_0^\ell) & L_0 \text{Var}_\ell(\hat{X}_0^\ell) L_0^\top + \text{Var}_\ell(\zeta_0) \end{bmatrix} \\ &\stackrel{(a)}{=} \begin{bmatrix} \text{Var}_\gamma(\hat{X}_0^\gamma) & \text{Cov}_\gamma(U_0^\gamma, \hat{X}_0^\gamma)^\top \\ \text{Cov}_\gamma(U_0^\gamma, \hat{X}_0^\gamma) & \text{Var}_\gamma(U_0^\gamma) \end{bmatrix} = \text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_0^\gamma \\ U_0^\gamma \end{bmatrix} \right), \end{aligned}$$

where (a) follows by substituting the expression for L_0 , and using the fact that $\text{Var}_\ell(\zeta_t) = \text{Var}_\gamma(U_t^\gamma) - \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma) \text{Var}_\gamma(\hat{X}_t^\gamma)^{-1} \text{Cov}_\gamma(U_t^\gamma, \hat{X}_t^\gamma)^\top$.

Let us now assume that $\text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_k^\gamma \\ U_k^\gamma \end{bmatrix} \right) = \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_k^\ell \\ U_k^\ell \end{bmatrix} \right)$ for some $t = k$. Therefore, at time $t = k+1$ we have

$$\begin{aligned} \text{Var}_\ell(\hat{X}_{k+1}^\ell) &= \text{Var}_\ell \left([A_k \ B_k] \begin{bmatrix} \hat{X}_k^\ell \\ U_k^\ell \end{bmatrix} + K_{k+1} \xi_{k+1} \right) \\ &= \text{Var}_\ell \left([A_k \ B_k] \begin{bmatrix} \hat{X}_k^\ell \\ U_k^\ell \end{bmatrix} \right) + \text{Var}_\ell(K_{k+1} \xi_{k+1}) \\ &= [A_k \ B_k] \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_k^\ell \\ U_k^\ell \end{bmatrix} \right) \begin{bmatrix} A_k^\top \\ B_k^\top \end{bmatrix} + \text{Var}_\ell(K_{k+1} \xi_{k+1}) \\ &\stackrel{(b)}{=} [A_k \ B_k] \text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_k^\gamma \\ U_k^\gamma \end{bmatrix} \right) \begin{bmatrix} A_k^\top \\ B_k^\top \end{bmatrix} + \text{Var}_\gamma(K_{k+1} \xi_{k+1}) \\ &= \text{Var}_\gamma(\hat{X}_{k+1}^\gamma), \end{aligned} \tag{20}$$

where (b) follows from the assumption that $\text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_k^\gamma \\ U_k^\gamma \end{bmatrix} \right) = \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_k^\ell \\ U_k^\ell \end{bmatrix} \right)$ and fact that the distributions of ξ_{k+1} under $\mathbb{P}_\ell$ and $\mathbb{P}_\gamma$ are

the same, since they depend only on the primitive random variables and not on the control law.

Furthermore, we may write

$$\begin{aligned} \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_{k+1}^\ell \\ U_{k+1}^\ell \end{bmatrix} \right) &= \begin{bmatrix} \text{Var}_\ell(\hat{X}_{k+1}^\ell) & \text{Var}_\ell(\hat{X}_{k+1}^\ell) L_{k+1}^\top \\ L_{k+1} \text{Var}_\ell(\hat{X}_{k+1}^\ell) & L_{k+1} \text{Var}_\ell(\hat{X}_{k+1}^\ell) L_{k+1}^\top \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \text{Var}_\ell(\zeta_{k+1}) \end{bmatrix} \\ &= \text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_{k+1}^\gamma \\ U_{k+1}^\gamma \end{bmatrix} \right), \end{aligned}$$

proving the claim that $\text{Var}_\gamma \left(\begin{bmatrix} \hat{X}_t^\gamma \\ U_t^\gamma \end{bmatrix} \right) = \text{Var}_\ell \left(\begin{bmatrix} \hat{X}_t^\ell \\ U_t^\ell \end{bmatrix} \right)$ for all t .

Step 3. From the above two steps, we have that

$$\begin{aligned} \mathbb{E}_\ell[\hat{Z}_t^\ell] &= [F_t \ G_t] \begin{bmatrix} \mathbb{E}_\ell[\hat{X}_t^\ell] \\ \mathbb{E}_\ell[U_t^\ell] \end{bmatrix} = [F_t \ G_t] \begin{bmatrix} \mathbb{E}_\gamma[\hat{X}_t^\gamma] \\ \mathbb{E}_\gamma[U_t^\gamma] \end{bmatrix} \\ &= \mathbb{E}_\gamma[\hat{Z}_t^\gamma], \end{aligned} \tag{21}$$

for all t , and similarly, one obtains that $\text{Var}_\ell(\hat{Z}_t^\ell) = \text{Var}_\gamma(\hat{Z}_t^\gamma)$.

Next, we may rewrite $\mathbb{E}_\ell[\|\hat{Z}_t^\ell - r_t\|_{S_t}^2]$ as

$$\begin{aligned} \mathbb{E}_\ell[\|\hat{Z}_t^\ell - r_t\|_{S_t}^2] &= \text{tr}(S_t \text{Var}_\ell(\hat{Z}_t^\ell)) + \|\mathbb{E}_\ell[\hat{Z}_t^\ell] - r_t\|_{S_t}^2 \\ &\stackrel{(a)}{=} \text{tr}(S_t \text{Var}_\gamma(\hat{Z}_t^\gamma)) + \|\mathbb{E}_\gamma[\hat{Z}_t^\gamma] - r_t\|_{S_t}^2 \\ &= \mathbb{E}_\gamma[\|\hat{Z}_t^\gamma - r_t\|_{S_t}^2], \end{aligned} \tag{22}$$

where (a) follows from the fact that $\text{Var}_\gamma(\hat{Z}_t^\gamma) = \text{Var}_\ell(\hat{Z}_t^\ell)$ and $\mathbb{E}_\gamma(\hat{Z}_t^\gamma) = \mathbb{E}_\ell(\hat{Z}_t^\ell)$ for all t . Similarly, one can verify that

$$\mathbb{E}_\ell[(\hat{Z}_T^\ell - \mu_f)(\hat{Z}_T^\ell - \mu_f)^\top] = \mathbb{E}_\gamma[(\hat{Z}_T^\gamma - \mu_f)(\hat{Z}_T^\gamma - \mu_f)^\top], \tag{23a}$$

$$\mathbb{E}_\ell[\hat{Z}_T^\ell] = \mathbb{E}_\gamma[\hat{Z}_T^\gamma]. \tag{23b}$$

Notice from (22) that, if the nonlinear control law γ satisfies the constraints $\mathbb{E}_\gamma[\|\hat{Z}_t^\gamma - r_t\|_{S_t}^2] \leq \bar{\delta}_t$, then the affine policy ℓ also satisfies the constraint $\mathbb{E}_\ell[\|\hat{Z}_t^\ell - r_t\|_{S_t}^2] \leq \bar{\delta}_t$. Similarly, from (23a) it follows that, if the policy γ satisfies the terminal constraints (11), then the policy ℓ also satisfies the constraints (11). Therefore, if γ is a feasible control policy for Problem 3, then ℓ is also a feasible policy for Problem 3. Furthermore, the cost under the policy ℓ is rewritten as

$$\begin{aligned} \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^{\ell\top} Q_t \hat{Z}_t^\ell \right] &= \sum_{t=0}^T \text{tr}(Q_t \text{Var}_\ell(\hat{Z}_t^\ell)) + \mathbb{E}[\hat{Z}_t^\ell]^\top Q_t \mathbb{E}[\hat{Z}_t^\ell] \\ &= \sum_{t=0}^T \text{tr}(Q_t \text{Var}_\gamma(\hat{Z}_t^\gamma)) + \mathbb{E}[\hat{Z}_t^\gamma]^\top Q_t \mathbb{E}[\hat{Z}_t^\gamma] \\ &= \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^{\gamma\top} Q_t \hat{Z}_t^\gamma \right], \end{aligned} \tag{24}$$

which proves that the cost under the policy ℓ is the same as the cost under the policy γ .

This analysis shows that it is sufficient to look for controllers of the form $U_t^\ell = L_t \hat{X}_t^\ell + \bar{u}_t + \zeta_t$. This affine policy is a randomized control law due to the presence of the random variable ζ_t . Using Lemma 3 one concludes that the performance of the control $U_t^\ell = L_t \hat{X}_t^\ell + \bar{u}_t + \mathbb{E}[\zeta_t]$ is no worse than the one using $U_t^\ell = L_t \hat{X}_t^\ell + \bar{u}_t + \zeta_t$. Therefore, it is sufficient to only consider the class of control policies of the form $U_t = L_t \hat{X}_t + \bar{u}_t$, where the matrix L_t and the vector $\bar{u}_t$ are deterministic. This completes the proof.

Having characterized the structure of an optimal control policy, in the next section we develop an optimization problem to find the optimal feedback gain L_t and the open loop input $\bar{u}_t$.

6. Optimal control via convex optimization

6.1. Augmented state representation

In the previous section, we showed that the class of affine controllers contains the optimal controller for Problem 3. In order to

⁶ For two random variables X, Y , we say $X = Y$ to imply the that $X = Y$ almost surely.

find the optimal affine controller, let us start by defining the augmented quantities as $\hat{Z} = [\hat{Z}_0^\top, \dots, \hat{Z}_T^\top]^\top$, $U = [U_0^\top, \dots, U_{T-1}^\top]^\top$, $\bar{u} = [\bar{u}_0^\top, \dots, \bar{u}_{T-1}^\top]^\top$, $\hat{X} = [\hat{X}_0^\top, \dots, \hat{X}_T^\top]^\top$ and $\Xi = [\xi_0^\top, \dots, \xi_T^\top]^\top$. Thus, the control law (16) can be written as

$$U = \begin{bmatrix} L_0 & 0 & \dots & 0 & 0 \\ 0 & L_1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & L_{T-1} & 0 \end{bmatrix} \hat{X} + \bar{u}.$$

Using the dynamics (8), the augmented state vector can be represented as⁷

$$\hat{X} = \mathcal{A}\hat{X} + \mathcal{B}\mathcal{L}\hat{X} + \mathcal{B}\bar{u} + \mathcal{K}\Xi,$$

where

$$\mathcal{L} = \begin{bmatrix} L_0 & 0 & \dots & 0 & 0 \\ 0 & L_1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & L_{T-1} & 0 \end{bmatrix},$$

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ A_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & A_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_{T-1} & 0 \end{bmatrix} \triangleq \text{subdiag}\{A_0, \dots, A_{T-1}\},$$

Similarly, $\mathcal{B} = \text{subdiag}\{B_0, \dots, B_{T-1}\}$ and $\mathcal{K} = \text{diag}\{K_0, \dots, K_T\}$. The augmented variables $\hat{X}$ and U are expressed as

$$\hat{X} = (I - (\mathcal{A} + \mathcal{B}\mathcal{L}))^{-1} (\mathcal{K}\Xi + \mathcal{B}\bar{u}), \quad (25a)$$

$$U = \mathcal{L} (I - (\mathcal{A} + \mathcal{B}\mathcal{L}))^{-1} (\mathcal{K}\Xi + \mathcal{B}\bar{u}) + \bar{u}. \quad (25b)$$

Notice that, due to the subdiagonal nature of $\mathcal{A} + \mathcal{B}\mathcal{L}$, the matrix $(I - (\mathcal{A} + \mathcal{B}\mathcal{L}))^{-1}$ always exists. Let us define the new variable $\mathcal{P} = \mathcal{L} (I - (\mathcal{A} + \mathcal{B}\mathcal{L}))^{-1}$ [31]. One can verify that (see Appendix A.3 for the proof) $(I - (\mathcal{A} + \mathcal{B}\mathcal{L}))^{-1} = (I - \mathcal{A})^{-1} (I + \mathcal{B}\mathcal{P})$, and $\mathcal{L} = (I + \mathcal{P}\mathcal{B})^{-1} \mathcal{P} (I - \mathcal{A})$. It then follows from (25) that,

$$\hat{X} = (I - \mathcal{A})^{-1} (I + \mathcal{B}\mathcal{P}) (\mathcal{K}\Xi + \mathcal{B}\bar{u}), \quad (26a)$$

$$U = \mathcal{P} (\mathcal{K}\Xi + \mathcal{B}\bar{u}) + \bar{u}, \quad (26b)$$

$$\begin{aligned} \hat{Z} &= \mathcal{F}\hat{X} + \mathcal{G}U \\ &= (\mathcal{F} (I - \mathcal{A})^{-1} (I + \mathcal{B}\mathcal{P}) + \mathcal{G}\mathcal{P}) (\mathcal{K}\Xi + \mathcal{B}\bar{u}) + \mathcal{G}\bar{u} \\ &= (\bar{\mathcal{F}} + (\mathcal{G} + \bar{\mathcal{F}}\mathcal{B})\mathcal{P}) \mathcal{K}\Xi + (\bar{\mathcal{F}}\mathcal{B} + (\mathcal{G} + \bar{\mathcal{F}}\mathcal{B})\mathcal{P}\mathcal{B} + \mathcal{G}) \bar{u} \\ &= (\bar{\mathcal{F}} + (\mathcal{G} + \bar{\mathcal{F}}\mathcal{B})\mathcal{P}) \mathcal{K}\Xi + (\bar{\mathcal{F}}\mathcal{B} + \mathcal{G})\bar{u} \\ &= \bar{\mathcal{P}}\mathcal{K}\Xi + \bar{\mathcal{G}}\bar{u} \end{aligned} \quad (26c)$$

$$(26d)$$

where $\mathcal{F} \triangleq \text{diag}\{F_0, \dots, F_T\}$, $\mathcal{G} \triangleq \text{diag}\{G_0, \dots, G_T\}$, and

$$\bar{\mathcal{F}} \triangleq \mathcal{F} (I - \mathcal{A})^{-1},$$

$$\bar{\mathcal{G}} \triangleq \bar{\mathcal{F}}\mathcal{B} + \mathcal{G},$$

$$\bar{\mathcal{P}} \triangleq \bar{\mathcal{F}} + \bar{\mathcal{G}}\mathcal{P},$$

$$v \triangleq (I + \mathcal{P}\mathcal{B})\bar{u}.$$

Since $\xi_t \perp \xi_s$ for all $t \neq s$, it follows that the matrix $\mathbb{E}[\Xi\Xi^\top]$ is block diagonal with diagonal entries $\mathbb{E}[\xi_t\xi_t^\top] = M_t$, where the matrix M_t is defined in Lemma 2.

Therefore, using (26d) one can rewrite $\hat{Z}_t = E_t\hat{Z}$ where $E_t = [0, \dots, 0, I, 0, \dots, 0] \in \mathbb{R}^{l \times l(T+1)}$, and

$$\begin{aligned} \mathbb{E} \left[\sum_{t=0}^T \hat{Z}_t^\top Q_t \hat{Z}_t \right] &= \mathbb{E}[\hat{Z}^\top Q \hat{Z}] = \text{tr} (Q \mathbb{E}[\hat{Z} \hat{Z}^\top]) \\ &= \text{tr} (Q \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top) + v^\top \bar{\mathcal{G}}^\top Q \bar{\mathcal{G}} v, \end{aligned}$$

where $Q = \text{diag}\{Q_0, \dots, Q_T\}$, and $\mathcal{M} = \text{diag}\{M_0, \dots, M_T\}$. Furthermore,

$$\begin{aligned} \mathbb{E}[\|\hat{Z}_t - r_t\|_{S_t}^2] &= \mathbb{E}[\hat{Z}_t^\top S_t \hat{Z}_t] - 2r_t^\top S_t \mathbb{E}[\hat{Z}_t] + r_t^\top S_t r_t \\ &= \text{tr} (E_t^\top S_t E_t \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top) + v^\top \bar{\mathcal{G}}^\top E_t^\top S_t E_t \bar{\mathcal{G}} v \\ &\quad - 2r_t^\top S_t E_t \bar{\mathcal{G}} v + r_t^\top S_t r_t \\ &= \text{tr} (E_t^\top S_t E_t \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top) + \|r_t - E_t \bar{\mathcal{G}} v\|_{S_t}^2. \end{aligned}$$

Therefore, the cost function and the constraints are quadratic in the optimization variables $\mathcal{P}$ and v . The new optimization problem is convex and it is formally presented in Problem 4.

Problem 4. Using the augmented variables, Problem 3 is equivalently rewritten as

$$\begin{aligned} \min_{\mathcal{P}, v} \quad & \text{tr} (Q \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top) + v^\top \bar{\mathcal{G}}^\top Q \bar{\mathcal{G}} v + c \\ \text{subject to} \quad & \|r_t - E_t \bar{\mathcal{G}} v\|_{S_t}^2 + \text{tr} (E_t^\top S_t E_t \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top) \leq \delta_t - \text{tr}(F_t^\top S_t F_t \Sigma_t), \\ & t = 0, \dots, T-1, \\ & E_T \bar{\mathcal{P}} \mathcal{K} \mathcal{M} \mathcal{K}^\top \bar{\mathcal{P}}^\top E_T^\top + (\mu_f - E_T \bar{\mathcal{G}} v)(\mu_f - E_T \bar{\mathcal{G}} v)^\top \leq \Sigma_f - F_T \Sigma_T F_T^\top, \\ & E_T \bar{\mathcal{G}} v = \mu_f, \quad \bar{\mathcal{P}} = \bar{\mathcal{F}} + \bar{\mathcal{G}}\mathcal{P}, \end{aligned}$$

where $c = \sum_{t=0}^T \text{tr}(Q_t F_t \Sigma_t F_t^\top)$.

Problem 4 can be solved using convex solvers, e.g., [32]. Once Problem 4 is solved, one can construct $\mathcal{L}$ and $\bar{u}$ from $\mathcal{P}$ and v using

$$\mathcal{L} = (I + \mathcal{P}\mathcal{B})^{-1} \mathcal{P} (I - \mathcal{A}), \quad \bar{u} = (I + \mathcal{P}\mathcal{B})^{-1} v.$$

6.2. Problem feasibility

We revisit Fact 1 and discuss the feasibility of the final optimization problem, and how it is influenced by controllability, observability, and the choice of the problem parameters such as r_t , S_t , δ_t , etc. Although controllability was not invoked in any of the preceding derivations, the final equality constraint in Problem 3, namely, $E_T \bar{\mathcal{G}} v = \mu_f$, requires that μ_f lies in the range space of $E_T \bar{\mathcal{G}}$. To ensure feasibility for an arbitrary μ_f , the matrix $E_T \bar{\mathcal{G}}$ must therefore be full row rank.

Upon closer examination, we observe that

$$E_T \bar{\mathcal{G}} = \underbrace{\left[F_T \prod_{k=1}^{T-1} A_k B_0 \quad F_T \prod_{k=2}^{T-1} A_k B_1 \quad \dots \quad F_T A_{T-1} B_{T-2} \quad (F_T B_{T-1} + G_T) \right]}_{\ell \times mT \text{ matrix}}.$$

In the special case where $G_T = 0$, a necessary condition for $E_T \bar{\mathcal{G}}$ to have full row rank is that the controllability matrix over horizon $T-1$ has full rank. That is,

$$\text{rank} \left(\begin{bmatrix} \prod_{k=1}^{T-1} A_k B_0 & \prod_{k=2}^{T-1} A_k B_1 & \dots & A_{T-1} B_{T-2} & B_{T-1} \end{bmatrix} \right) = n.$$

In addition to this controllability condition, if F_T has full row rank, then it is sufficient to ensure that $E_T \bar{\mathcal{G}}$ has full row rank, and thus the equality constraint can be satisfied for arbitrary μ_f .

Observability affects feasibility in a more nuanced way. Even if the system is observable, the problem may still be infeasible. It is not simply a binary question of whether the system is observable, but rather *how well* it is observed. In particular, the estimation error covariance Σ_t influences the feasibility of the inequality constraints through the terms $\text{tr}(F_t^\top S_t F_t \Sigma_t)$ and $F_T \Sigma_T F_T^\top$. Systems with smaller Σ_t generally exhibit better feasibility. However, F_t and S_t also play critical roles. Even when Σ_t has relatively large values, the quantities $\text{tr}(F_t^\top S_t F_t \Sigma_t)$ and $F_T \Sigma_T F_T^\top$ can still be small for certain choices of F_t and S_t , leading to improved feasibility.

If a problem turns out to be infeasible, identifying the source of the infeasibility, i.e., the specific constraints that are violated, can be helpful for restoring feasibility. For instance, if the infeasibility arises from the inequality constraints, it can often be addressed by increasing δ_t or by reducing the eigenvalues of S_t (e.g., by setting $S'_t = \alpha S_t$ for some $\alpha \in (0, 1)$).

⁷ For simplicity we assume $\mu_0 = 0$.

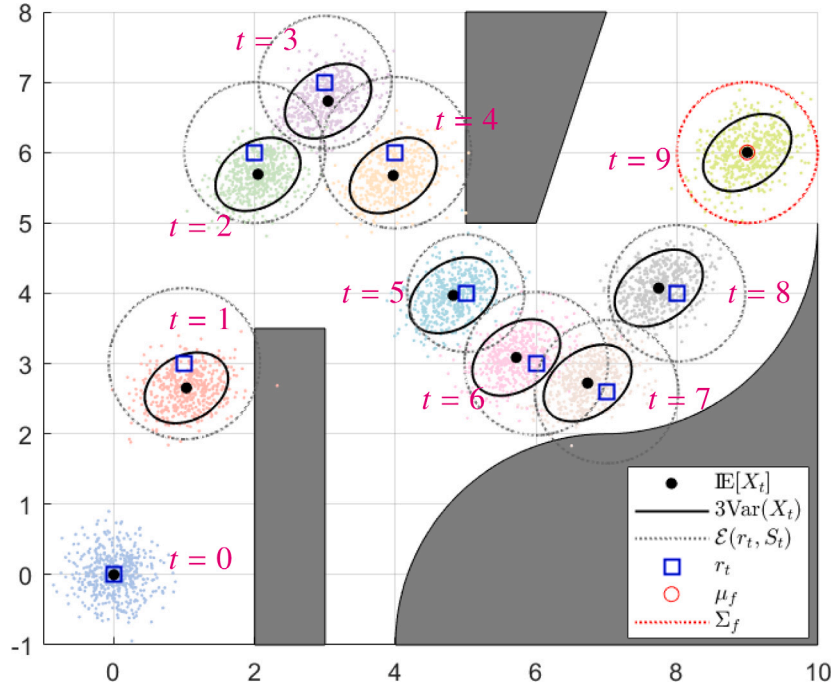


Fig. 1. Trajectories of the system states and the corresponding covariance ellipses. Results from 500 random trials are shown as scattered colored dots.

7. Numerical example

We illustrate our theory on an output-feedback covariance steering problem in the 2D environment shown in Fig. 1. This example is motivated by a robotic path planning problem in the presence of obstacles. The dark gray regions represent unsafe areas, and therefore the state X_t must remain outside these regions at all times with high probability, as enforced by the chance constraints in (2b). The system dynamics, sensor measurements, and output variables are defined as:

$$X_{t+1} = \begin{bmatrix} -0.4 & 1.2 \\ 0.5 & 0.3 \end{bmatrix} X_t + U_t + W_t;$$

$$Y_t = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} X_t + V_t;$$

$$Z_t = \begin{cases} \begin{bmatrix} I_2 \\ 0_2 \end{bmatrix} X_t + \begin{bmatrix} 0_2 \\ I_2 \end{bmatrix} U_t & t \neq T, \\ \begin{bmatrix} X_t \\ 0_2 \end{bmatrix}, & t = T, \end{cases}$$

where I_n and 0_n denote the identity and zero matrices of dimension $n \times n$. The process and measurement noises are distributed as $W_t \sim \mathcal{N}(0, 0.05I_2)$ and $V_t \sim \mathcal{N}(0, 0.05I_2)$, respectively, and the initial state is distributed as $X_0 \sim \mathcal{N}(0, 0.1I_2)$. We solve the optimization problem using the YALMIP toolbox [33] with the SDPT3 solver [34]. The other parameters of Problem 1 are chosen as follows.

$$\begin{aligned} T &= 9, \quad Q_t = \frac{1}{2}I_4, \quad \delta_t = 0.3, \quad S_t = \text{diag}\{\epsilon_t I_2, 0_2\}, \\ \epsilon_{0:9} &= [1.16, 0.93, 1.0, 1.06, 0.93, 1.20, 0.98, 0.98, 1.03, 1.17] \\ r_{0:9} &= \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 0 & 3 & 6 & 7 & 6 & 4 & 3 & 2.6 & 4 & 6 \end{bmatrix}, \\ \mu_f &= [9, 6, 0, 0]^T, \text{ and } \Sigma_f = \begin{bmatrix} I_2 & 0_2 \\ 0_2 & 0_2 \end{bmatrix}. \end{aligned}$$

As discussed earlier, the chance constraints are approximated as ellipsoidal constraints. Furthermore, in this example, we consider chance constraints only on the states and not on the control inputs. Hence, we choose S_t to be of the form $S_t = \text{diag}\{\epsilon_t I_2, 0_2\}$, where ϵ_t controls the tightness of the constraint (i.e., the larger the ϵ_t , the tighter the constraint). Given the structure of S_t , the (density-independent) relaxed

chance constraints (5b) become $\mathbb{E}[\|X_t - r_t\|^2] \leq \delta_t/\epsilon_t$. The chance constraint regions $\mathcal{E}_t(r_t, S_t) = \{x \in \mathbb{R}^2 \mid \epsilon_t \|x - r_t\|^2 \leq 1\}$ are shown using dotted gray circles in Fig. 1.

In these chance constraints, r_t represents the *reference point* at time t , which is plotted using empty blue squares in Fig. 1. We pick r_t such that they are reasonably away from the unsafe regions. The chance constraints are designed to keep the state X_t close to r_t . Note that we do not enforce the mean of the state, $\mathbb{E}[X_t]$, to equal r_t ; instead, the chance constraint ensures that the probability of deviation from the reference trajectory remains below the threshold δ_t .

The mean of the state, $\mathbb{E}[X_t]$, at each time t is plotted using filled black circles, and the solid ellipses represent the scaled (three times) state covariances under the designed controller. We also conducted 500 random trials, and the resulting states are shown as scattered dots. From these random trials, we numerically computed the probability of satisfying the chance constraints, i.e., $\mathbb{P}(X_t \in \mathcal{E}_t(r_t, S_t))$; see Table 1.

These results demonstrate that the chance constraints are satisfied with high probability, as expected from the theory.

7.1. Experiment with uniform noise

Next, we consider the same setup as in the previous experiment, except now the initial state X_0 and process noise W_t are drawn from a *uniform distribution*: $[X_0]_i \sim \mathcal{U}(-\sqrt{0.3}, \sqrt{0.3})$ and $[W_t]_i \sim \mathcal{U}(-\sqrt{0.15}, \sqrt{0.15})$, where $[v]_i$ for $i = 1, 2$ denotes the i th component of vector v . We purposefully chose the support of the X_0 (respectively, W_t) distribution to be $[-\sqrt{0.3}, \sqrt{0.3}]$ (respectively, $[-\sqrt{0.15}, \sqrt{0.15}]$) so that the mean and covariance of X_0 (resp. W_t) match those of the previous experiment with Gaussian noise (i.e., $\mathbb{E}[X_0] = \mathbb{E}[W_t] = 0$ and $\text{Var}(X_0) = 0.1I_2, \text{Var}(W_t) = 0.05I_2$). The uniform distribution has the maximum entropy among all distributions with bounded support and given second moment, thereby highlighting the efficacy of our approach under a distribution shift.

It should be noted that the Kalman filter is no longer optimal under uniform noise. Nonetheless, the output feedback covariance controller still performs well, as demonstrated in Fig. 2. The $3\text{Var}(X_t)$ ellipses are the same as those shown in Fig. 1. Results from 500 random trials are also shown in Fig. 2 for comparison with Fig. 1.

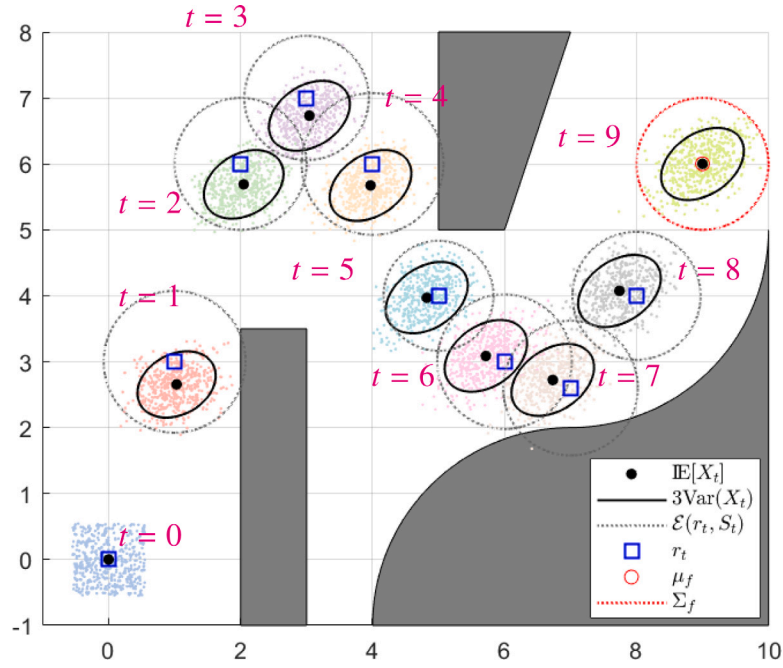


Fig. 2. Trajectories of the system states and covariance ellipses under uniformly distributed initial state and dynamics noise. Results from 500 random trials are shown as scattered colored dots.

Table 1

Empirical computation of $\mathbb{P}(X_t \in \mathcal{E}(r_t, S_t))$ from 500 random trials under Gaussian noise.

$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$
0.998	0.980	0.976	0.972	0.974	0.970	0.964	0.968	0.970	0.962

Table 2

Empirical computation of $\mathbb{P}(X_t \in \mathcal{E}(r_t, S_t))$ from 500 random trials under uniform noise.

$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$
1.000	0.984	0.974	0.976	0.966	0.968	0.970	0.984	0.970	0.974

Overall, the performance remains similar (see [Tables 1](#) and [2](#)), with only minor differences. For example, in [Fig. 1](#), one of the trials hits the rectangular obstacle near $x = 2$ at $t = 1$, which does not occur under uniform noise. In both experiments, one of the trials hits the curved obstacle at time $t = 7$; however, under uniform noise, it penetrates slightly deeper into the obstacle.

8. Conclusion

In this work, we considered the chance-constrained optimal covariance steering control problem for linear systems with partial and noisy observations. We approximate the chance constraint using a tractable formulation and characterize the optimal controller for the approximate problem. This results in a feasibility-preserving *sufficient approximation* in the sense that any feasible solution of the approximate problem will satisfy all the constraints of the original problem. We first prove that it is sufficient to consider the class of affine controllers when designing the optimal controller. Based on this observation, we propose a convex optimization problem to construct the optimal controller.

The work relies on the *distribution-independent approximation* of the chance constraint using the Markov inequality. Such inequalities can be, however, conservative. Therefore, designing a less conservative and distribution-independent approximation of the chance constraints is a

potential direction for future study. We emphasize that the use of Markov's inequality introduces and preserves problem convexity, which is instrumental in formulating the convex optimization problem in [Problem 4](#) and deriving our main result in [Theorem 1](#). In contrast, other forms of probabilistic inequalities may not inherently preserve such desirable properties, and would therefore require additional approximations to achieve a similar convex formulation. We also note that the conservativeness of Markov's inequality can be mitigated by reformulating the chance constraint $\mathbb{P}[\|Z_t - r_t\|_{S_t}^2 > 1] \leq \delta_t$ as $\mathbb{E}[\|Z_t - r_t\|_{S_t}^2] \leq \alpha \delta_t$ for some $\alpha > 1$. In this case, α must be chosen judiciously to ensure that the Markov bound does not become an under-approximation of the original chance constraint.

Finally, although the convexity of the optimization problem enables us to use efficient algorithms to find an optimal solution, however, the size of the problem can quickly become large. The complexity of solving [Problem 4](#) grows quadratically with the terminal time T (P matrix has $O(T^2)$ block elements). Therefore, a sequential receding horizon optimization framework instead of the augmented state representation is desired for problems with larger time horizons.

CRedit authorship contribution statement

Dipankar Maity: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Panagiotis Tsiotras:** Writing – review & editing, Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Support for this work has been provided by NSF award 2008686.

Appendix A

A.1. Proof of Lemma 1

It is sufficient to show that $\{Y_{0:t}, U_{0:t}\}$ and $\{\xi_{0:t}, U_{0:t}\}$ can be constructed from each other. We write $\{Y_{0:t}, U_{0:t}\} \equiv \{\xi_{0:t}, U_{0:t}\}$ to denote that these two sets of random variables are equivalent, i.e., one can be constructed from the other.

Let us consider a state-process X_t^{new} and an observation-process Y_t^{new} as follows

$$X_t^{\text{new}} = X_t - \sum_{k=0}^{t-1} \prod_{\ell=k+1}^{t-1} A_\ell B_k U_k - \prod_{k=0}^{t-1} A_k \mu_0, \quad (\text{A.1a})$$

$$Y_t^{\text{new}} = C_t X_t^{\text{new}} + V_t. \quad (\text{A.1b})$$

We use the convention that for all $b < a$, $\sum_a^b(\cdot) = 0$ and $\prod_a^b(\cdot) = 1$. Therefore, it follows that

$$X_{t+1}^{\text{new}} = A_t X_t^{\text{new}} + W_t, \quad (\text{A.2a})$$

$$Y_t^{\text{new}} = C_t X_t^{\text{new}} + V_t, \quad (\text{A.2b})$$

$$X_0^{\text{new}} = X_0 - \mu_0 = W_{-1} \sim \mathcal{N}(0, \Sigma_0). \quad (\text{A.2c})$$

Here X_t^{new} is the uncontrolled process associated with X_t , which is independent of the control strategy. Using this definition of X_t^{new} and Y_t^{new} , we have $X_t = X_t^{\text{new}} + \varphi(t, U_{0:t-1})$ and $Y_t = Y_t^{\text{new}} + C_t \varphi(t, U_{0:t-1})$ where $\varphi(t, U_{0:t-1}) = \sum_{k=0}^{t-1} \prod_{\ell=k+1}^{t-1} A_\ell B_k U_k + \prod_{k=0}^{t-1} A_k \mu_0$. Therefore, we have the equivalence between the information sets $\{Y_{0:t-1}, U_{0:t-1}\}$ and $\{Y_{0:t-1}^{\text{new}}, U_{0:t-1}\}$, i.e., one can be constructed from the other.

The innovation process associated with system (A.2) is given by

$$\xi_t^{\text{new}} = Y_t^{\text{new}} - \mathbb{E}[Y_t^{\text{new}} | Y_{0:t-1}^{\text{new}}].$$

It is a standard result in Kalman filtering theory that $\{Y_{0:t}^{\text{new}}\} \equiv \{\xi_{0:t}^{\text{new}}\}$. Since ξ_t is the innovation process associated with the system (1), it can be shown that $\xi_t^{\text{new}} = \xi_t$ for all $t \geq 0$. In order to prove $\xi_t^{\text{new}} = \xi_t$, notice that

$$\begin{aligned} \xi_t &= Y_t - \mathbb{E}[Y_t | Y_{0:t-1}, U_{0:t-1}] \\ &= Y_t^{\text{new}} + C_t \varphi(t, U_{0:t-1}) - \mathbb{E}[Y_t^{\text{new}} | Y_{0:t-1}, U_{0:t-1}] \\ &\quad - \mathbb{E}[C_t \varphi(t, U_{0:t-1}) | Y_{0:t-1}, U_{0:t-1}] \\ &= Y_t^{\text{new}} - \mathbb{E}[Y_t^{\text{new}} | Y_{0:t-1}^{\text{new}}, U_{0:t-1}] \\ &= Y_t^{\text{new}} - \mathbb{E}[Y_t^{\text{new}} | Y_{0:t-1}^{\text{new}}] = \xi_t^{\text{new}}. \end{aligned}$$

Thus, $\{\xi_{0:t}, U_{0:t}\} \equiv \{\xi_{0:t}^{\text{new}}, U_{0:t}\} \equiv \{Y_{0:t}^{\text{new}}, U_{0:t}\} \equiv \{Y_{0:t}, U_{0:t}\}$, and the Lemma is proved. ■

A.2. Proof of Lemma 2

The standard results of Kalman filtering theory hold for the process X_t^{new} with observation Y_t^{new} defined in (A.2). It follows that $\xi_{0:T}^{\text{new}}$ is a sequence of uncorrelated Gaussian noises. Thus, using standard Kalman filtering theory, we define

$$e_t^{\text{new}} = X_t^{\text{new}} - \mathbb{E}[X_t^{\text{new}} | Y_{0:t-1}^{\text{new}}], \quad (\text{A.3a})$$

$$\Delta_t^{\text{new}} = X_t^{\text{new}} - \mathbb{E}[X_t^{\text{new}} | Y_{0:t}^{\text{new}}], \quad (\text{A.3b})$$

$$\Sigma_{t|t-1} = \mathbb{E}[e_t^{\text{new}} e_t^{\text{new}T}], \quad (\text{A.3c})$$

$$\Sigma_t = \mathbb{E}[\Delta_t^{\text{new}} \Delta_t^{\text{new}T}]. \quad (\text{A.3d})$$

Moreover,

$$\mathbb{E}[X_t^{\text{new}} | Y_{0:t}^{\text{new}}] = \mathbb{E}[X_t^{\text{new}} | Y_{0:t-1}^{\text{new}}] + K_t \xi_t^{\text{new}},$$

where $K_t = \Sigma_{t|t-1} C_t^T (C_t \Sigma_{t|t-1} C_t^T + \mathcal{V}_t)^{-1}$ is the Kalman gain at time t . Thus, $\Delta_t^{\text{new}} = e_t^{\text{new}} - K_t \xi_t^{\text{new}} = (I - K_t C_t) e_t^{\text{new}} - K_t V_t$. The initial conditions

are $e_0^{\text{new}} = X_0^{\text{new}} \sim \mathcal{N}(0, \Sigma_0)$ and $\Sigma_{0|-1} = \Sigma_0$. Therefore, $\mathbb{E}[\xi_t^{\text{new}}] = 0$ and, $M_t, \Sigma_{t|t-1}$ and Σ_t satisfy

$$\begin{aligned} M_t &= \mathbb{E}[(C_t e_t^{\text{new}} + V_t)(C_t e_t^{\text{new}} + V_t)^T] \\ &= C_t \Sigma_{t|t-1} C_t^T + \mathcal{V}_t, \\ \Sigma_{t|t-1} &= \mathbb{E}[e_t^{\text{new}} e_t^{\text{new}T}] \\ &= \mathbb{E}[(A_{t-1} \Delta_{t-1}^{\text{new}} + W_{t-1})(A_{t-1} \Delta_{t-1}^{\text{new}} + W_{t-1})^T] \\ &= A_{t-1} \Sigma_{t-1} A_{t-1}^T + \mathcal{W}_{t-1}, \\ \Sigma_t &= \mathbb{E}[(I - K_t C_t) e_t^{\text{new}} e_t^{\text{new}T} (I - K_t C_t)^T + K_t \mathcal{V}_t K_t^T] \\ &= (I - K_t C_t) \Sigma_{t|t-1} (I - K_t C_t)^T + K_t \mathcal{V}_t K_t^T \\ &= \Sigma_{t|t-1} - \Sigma_{t|t-1} C_t^T M_t^{-1} C_t \Sigma_{t|t-1}, \end{aligned}$$

where $K_t = \Sigma_{t|t-1} C_t^T M_t^{-1}$ is the Kalman gain. In order to prove that $\{\xi_0, \dots, \xi_t\}$ is a sequence of uncorrelated random variables, we show that, for all $t > s$,

$$\begin{aligned} \mathbb{E}[\xi_t \xi_s^T] &\stackrel{(a)}{=} \mathbb{E}[\mathbb{E}[\xi_t \xi_s^T | Y_{0:s}, U_{0:s}]] \stackrel{(b)}{=} \mathbb{E}[\mathbb{E}[\xi_t | Y_{0:s}, U_{0:s}] \xi_s^T] \\ &= \mathbb{E}[\mathbb{E}[\xi_t] \xi_s^T] \stackrel{(c)}{=} 0 = \mathbb{E}[\xi_t] \mathbb{E}[\xi_s^T]^T, \end{aligned}$$

where (a) follows from the law of iterated expectations, (b) follows from the fact ξ_s is $\{Y_{0:s}, U_{0:s}\}$ measurable, and (c) follows from the fact that ξ_t is independent of $\{Y_{0:s}, U_{0:s}\}$ for all $s < t$. This concludes the proof. ■

A.3. Relationship between $\mathcal{L}$ and $\mathcal{P}$

Lemma 4. Let $\mathcal{P} = \mathcal{L}(I - (A + B\mathcal{L}))^{-1}$. Then,

- (i) $(I - (A + B\mathcal{L}))^{-1} = (I - A)^{-1}(I + B\mathcal{P})$,
- (ii) $\mathcal{L} = (I + \mathcal{P}B)^{-1}\mathcal{P}(I - A)$.

Proof. Notice that since $A, B, \mathcal{L}$ are lower-triangular, the matrix $\mathcal{P}$ is lower-triangular. It follows that,

$$\begin{aligned} \mathcal{P} &= \mathcal{L}(I - (A + B\mathcal{L}))^{-1} \\ \implies B\mathcal{P} &= B\mathcal{L}(I - (A + B\mathcal{L}))^{-1} \\ &= -I + (I - A)(I - (A + B\mathcal{L}))^{-1} \\ \implies (I - A)^{-1}(I + B\mathcal{P}) &= (I - (A + B\mathcal{L}))^{-1} \end{aligned}$$

Therefore, statement (i) of the lemma is proved. In order to prove statement (ii), we proceed as follows

$$\begin{aligned} \mathcal{P} &= \mathcal{L}(I - (A + B\mathcal{L}))^{-1} \\ \implies \mathcal{P}(I - (A + B\mathcal{L})) &= \mathcal{L} \\ \implies \mathcal{P}(I - A) &= (I + \mathcal{P}B)\mathcal{L} \\ \implies \mathcal{L} &= (I + \mathcal{P}B)^{-1}\mathcal{P}(I - A). \end{aligned}$$

Due to the sub-diagonal structure of B , the matrix $(I + \mathcal{P}B)$ is invertible. This proves statement (ii). ■

Data availability

Data will be made available on request.

References

- [1] E.G. Collins, R.E. Skelton, Covariance control discrete systems, in: Conference on Decision and Control, pp. 542–547.
- [2] A. Hotz, R.E. Skelton, Covariance control theory, Internat. J. Control 46 (1) (1987) 13–32.
- [3] Y. Chen, T.T. Georgiou, M. Pavon, Optimal steering of a linear stochastic system to a final probability distribution, Part I, IEEE Trans. Autom. Control 61 (5) (2015) 1158–1169.
- [4] Y. Chen, T.T. Georgiou, M. Pavon, Optimal steering of a linear stochastic system to a final probability distribution, Part II, IEEE Trans. Autom. Control 61 (5) (2015) 1170–1180.

- [5] A. Halder, E.D. Wendel, Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost, in: American Control Conference, 2016, pp. 7249–7254.
- [6] Y. Chen, T.T. Georgiou, M. Pavon, Optimal steering of a linear stochastic system to a final probability distribution—Part III, *IEEE Trans. Autom. Control* 63 (9) (2018) 3112–3118.
- [7] Y. Chen, T.T. Georgiou, M. Pavon, On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint, *J. Optim. Theory Appl.* 169 (2) (2016) 671–691.
- [8] K.F. Caluya, A. Halder, Reflected Schrödinger bridge: Density control with path constraints, in: American Control Conference, 2021, pp. 1137–1142.
- [9] T. Iwasaki, R.E. Skelton, Quadratic optimization for fixed order linear controllers via covariance control, in: American Control Conference, 1992, pp. 2866–2870.
- [10] J.-H. Xu, R.E. Skelton, An improved covariance assignment theory for discrete systems, *IEEE Trans. Autom. Control* 37 (10) (1992) 1588–1591.
- [11] K. Yasuda, R.E. Skelton, K.M. Grigoriadis, Covariance controllers: A new parametrization of the class of all stabilizing controllers, *Automatica* 29 (3) (1993) 785–788.
- [12] K.M. Grigoriadis, R.E. Skelton, Minimum-energy covariance controllers, *Automatica* 33 (4) (1997) 569–578.
- [13] E. Bakolas, Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints, *Automatica* 91 (2018) 61–68.
- [14] E. Bakolas, Optimal covariance control for discrete-time stochastic linear systems subject to constraints, in: Conference on Decision and Control, 2016, pp. 1153–1158.
- [15] M. Goldshtein, P. Tsiotras, Finite-horizon covariance control of linear time-varying systems, in: Conference on Decision and Control, 2017, pp. 3606–3611.
- [16] K. Okamoto, M. Goldshtein, P. Tsiotras, Optimal covariance control for stochastic systems under chance constraints, *IEEE Control. Syst. Lett.* 2 (2) (2018) 266–271.
- [17] L. Blackmore, M. Ono, Convex chance constrained predictive control without sampling, in: AIAA Guidance, Navigation, and Control Conference, 2009, p. 5876.
- [18] A. Geletu, M. Klöppel, H. Zhang, P. Li, Advances and applications of chance-constrained approaches to systems optimisation under uncertainty, *Int. J. Syst. Sci.* 44 (7) (2013) 1209–1232.
- [19] M. Farina, L. Giulioni, R. Scattolini, Stochastic linear model predictive control with chance constraints—a review, *J. Process Control* 44 (2016) 53–67.
- [20] A. Mesbah, Stochastic model predictive control: An overview and perspectives for future research, *IEEE Control Syst. Mag.* 36 (6) (2016) 30–44.
- [21] L. Blackmore, M. Ono, B.C. Williams, Chance-constrained optimal path planning with obstacles, *IEEE Trans. Robot.* 27 (6) (2011) 1080–1094.
- [22] K. Okamoto, P. Tsiotras, Optimal stochastic vehicle path planning using covariance steering, *IEEE Robot. Autom. Lett.* 4 (3) (2019) 2276–2281.
- [23] J. Ridderhof, P. Tsiotras, Minimum-fuel powered descent in the presence of random disturbances, in: AIAA Scitech Forum, 2019, p. 0646.
- [24] K. Okamoto, P. Tsiotras, Input hard constrained optimal covariance steering, in: Conference on Decision and Control, 2019, pp. 3497–3502.
- [25] J. Ridderhof, K. Okamoto, P. Tsiotras, Nonlinear uncertainty control with iterative covariance steering, in: Conference on Decision and Control, 2019, pp. 3484–3490.
- [26] I.M. Balci, E. Bakolas, B. Vlahov, E. Theodorou, Constrained covariance steering based tube-MPPI, in: American Control Conference, Atlanta, GA, 2022, pp. 4197–4202.
- [27] E. Bakolas, Covariance control for discrete-time stochastic linear systems with incomplete state information, in: American Control Conference, 2017, pp. 432–437.
- [28] M. Farina, L. Giulioni, L. Magni, R. Scattolini, An approach to output-feedback MPC of stochastic linear discrete-time systems, *Automatica* 55 (2015) 140–149.
- [29] J. Ridderhof, K. Okamoto, P. Tsiotras, Chance constrained covariance control for linear stochastic systems with output feedback, in: Conference on Decision and Control, 2020, pp. 1758–1763.
- [30] D. Maity, P. Tsiotras, Optimal controller and quantizer selection for partially observable linear-quadratic-Gaussian systems, *SIAM J. Control Optim.* 61 (4) (2023) 2682–2707.
- [31] J. Skaf, S.P. Boyd, Design of affine controllers via convex optimization, *IEEE Trans. Autom. Control* 55 (11) (2010) 2476–2487.
- [32] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014.
- [33] J. Lofberg, YALMIP: A toolbox for modeling and optimization in MATLAB, in: International Conference on Robotics and Automation, 2004, pp. 284–289.
- [34] K.-C. Toh, M.J. Todd, R.H. Tütüncü, SDPT3—a MATLAB software package for semidefinite programming, version 1.3, *Optim. Methods Softw.* 11 (1–4) (1999) 545–581.