

# Stochastic LQR Design With Disturbance Preview

Jietian Liu

*Electrical Eng. and Computer Science  
University of Michigan  
jietian@umich.edu*

Laurent Lessard

*Mechanical and Industrial Eng.  
Northeastern University  
l.lessard@northeastern.edu*

Peter Seiler

*Electrical Eng. and Computer Science  
University of Michigan  
pseiler@umich.edu*

**Abstract**—This paper considers the discrete-time, stochastic LQR problem with  $p$  steps of disturbance preview information where  $p$  is finite. We first derive the solution for this problem on a finite horizon with linear, time-varying dynamics and time-varying costs. Next, we derive the solution on the infinite horizon with linear, time-invariant dynamics and time-invariant costs. Our proofs rely on the well-known principle of optimality. We provide an independent proof for the principle of optimality that relies only on nested information structure. Finally, we show that the finite preview controller converges to the optimal noncausal controller as the preview horizon  $p$  tends to infinity. We also provide a simple example to illustrate both the finite and infinite horizon results.

**Index Terms**—Stochastic LQR, discrete time, disturbance preview

## I. INTRODUCTION

This paper considers control design for a discrete-time, stochastic Linear Quadratic Regulator (LQR) problem with finite preview information of the disturbance. We consider both the finite and infinite horizon problems. Our specific formulation is motivated by recent work on regret-based control [1], [2], [3], [4], [5], [6]. These regret-based formulations often use a controller with full knowledge of the disturbance as a baseline for performance. The optimal controller with finite preview, as derived in our paper, can provide an alternative baseline for regret-based formulations.

There is a large literature on finite preview control for stochastic LQR/LQG and its closely related  $H_2$  formulation. These results exist in both continuous-time [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17] and discrete-time [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]. A common approach in discrete-time is to construct an augmented system with a chain of delays to store the disturbance preview information. This technique simplifies the preview control design as the standard LQR/LQG/ $H_2$  solutions can be applied to this augmented model [9], [18], [19], [11], [12], [21], [22], [23], [24], [14], [25], [15], [16], [26], [27]. The work in [13] is an exception as it uses calculus of variations to directly derive necessary conditions for optimality.

Our paper makes the following contributions to this literature. First, we solve the stochastic LQR problems with finite disturbance preview on both finite and infinite time horizons (Sections V-A and V-B). In principle, these solutions follow from existing stochastic LQR results by using the state augmentation procedure. However, we provide an independent proof for the finite-horizon and infinite-horizon problems

based on the principle of optimality. Moreover, we provide a novel proof for the principle of optimality in this setting (Section IV). Our proof only relies on the assumption of nested information and may be of independent interest. Finally, we compare the performance of the finite disturbance preview controller against the optimal noncausal controller (Theorem 11.2.1 of [28]). The optimal noncausal controller has full knowledge of the future disturbance. We show that the finite preview controller converges, in a certain sense, to the optimal noncausal controller as the preview horizon tends to infinity (Section V-C). Finally, we illustrate the results with a simple example (Section VI).

Two of the most related works are [13] and [23]. Calculus of variations is used in [13] to derive a deterministic preview control solution. The approach starts by deriving the optimal non-causal controller with full knowledge of future disturbance, similar to that given in Theorem 11.2.1 of [28]. This optimal non-causal controller is then converted to a finite-preview controller by taking the expectation for unavailable future (post-preview) information which is zero. In [23], an LQG problem is solved with a cost function that includes cross terms between the state and input. They address the finite preview problem using an augmented state approach. The results in both [13] and [23] are similar to those contained in our paper. The key difference lies in the approach used to solve the problem. Specifically, our proof directly uses a version of the principle of optimality with nested information. This eliminates the need for state augmentation and hence our optimal controller is directly expressed with dynamics of the same order of the plant (not including augmented states). Moreover, our finite horizon derivation can directly accommodate non-zero initial conditions.

## II. NOTATION

Let  $\mathbb{R}^n$  and  $\mathbb{R}^{n \times m}$  denote the sets of real  $n \times 1$  vectors and  $n \times m$  matrices, respectively. The superscript  $\top$  denotes the matrix transpose. Moreover, if  $M \in \mathbb{R}^{n \times n}$  then  $M \succ 0$  denotes that  $M$  is a symmetric positive definite matrix. Similarly, we use  $\succeq$ ,  $\prec$ , and  $\preceq$  for positive semidefinite, negative definite, and negative semidefinite matrices, respectively. A matrix  $A \in \mathbb{R}^{n \times n}$  is said to be Schur if the spectral radius is  $< 1$ . We also introduce random variables in the problem formulation and use  $\mathbb{E}$  to denote the expectation. Finally, we denote the Kronecker delta  $\delta_{ij}$  where  $\delta_{ij} = 1$  if  $i = j$  and  $\delta_{ij} = 0$  otherwise.

### III. PROBLEM FORMULATION

We formulate the stochastic Linear Quadratic Regulator (LQR) problem with preview information of the disturbance. The problem is stated on both finite and infinite horizons.

Consider the following discrete-time, linear time-varying (LTV) plant  $P$  with the following state-space representation:

$$x_{t+1} = A_t x_t + B_{u,t} u_t + B_{w,t} w_t, \quad (1)$$

where  $x_t \in \mathbb{R}^{n_x}$ ,  $u_t \in \mathbb{R}^{n_u}$ , and  $w_t \in \mathbb{R}^{n_w}$  are the state, input, and disturbance at time  $t$ , respectively. The state matrices have compatible dimensions, e.g.  $A_t \in \mathbb{R}^{n_x \times n_x}$ , and are defined on a finite horizon  $t = 0, 1, \dots, T-1$  with  $T < \infty$ . We assume the disturbance is the independent and identically distributed (IID) noise:  $\mathbb{E}[w_i] = 0$  and  $\mathbb{E}[w_i w_j^\top] = \delta_{ij} I$  for  $i, j \in \{0, 1, \dots, T-1\}$ . We allow for the initial condition  $x_0$  to be non-zero.

The goal is to design a controller to reject the effect of the disturbance. We assume the controller has measurements of the state and  $p$  steps of disturbance preview where  $p \geq 0$ . Specifically, the controller has access to the following information at time  $t$ :

$$\begin{aligned} i_t &:- \{x_0, \dots, x_t, w_0, \dots, w_{t+p}\} \text{ if } t \leq T-p-1 \\ i_t &:- \{x_0, \dots, x_t, w_0, \dots, w_{T-1}\} \text{ otherwise.} \end{aligned} \quad (2)$$

The available information depends on the amount of preview  $p \geq 0$ . Also note that the information  $i_t$  contains some redundant information. However, it is important that this information pattern is nested, i.e.  $i_0 \subseteq i_1 \subseteq \dots \subseteq i_{T-1}$ . This allows the controller, in general, to access past information as time goes on. This nesting structure is also used later in our formal proofs.

We allow for controllers that can take random actions based on the available information. Thus the controller policy at time  $t$  is specified by a probability density function over possible actions given the available information:

$$K_t(u_t, i_t) := \text{Prob}(u_t | i_t) \quad (3)$$

This general formulation includes deterministic policies as a special case, which corresponds to functions of the form  $u_t = k_t(i_t)$ . We'll denote the sequence of control policies by  $K := \{K_0, \dots, K_{T-1}\}$ . The performance is evaluated with a linear quadratic cost. In particular, the cost for a sequence of policies  $K$  evaluated on a specific disturbance sequence  $w$  and initial condition  $x_0$  is defined as follows:

$$J_T(K, w, x_0) := x_T^\top Q_T x_T + \sum_{t=0}^{T-1} x_t^\top Q_t x_t + u_t^\top R_t u_t, \quad (4)$$

where  $Q_t \succeq 0$  for  $t = 0, \dots, T$  and  $R_t \succ 0$  for  $t = 0, \dots, T-1$ . These matrices define the state and control costs.

We now state the finite-horizon (FH), stochastic LQR problem with preview information.

**Problem 1 (FH Stochastic LQR With Preview).** Consider the LTV plant (1) defined on  $t = 0, \dots, T-1$  and cost function in (4) with  $Q_t \succeq 0$  for  $t = 0, \dots, T$  and  $R_t \succ 0$

for  $t = 0, \dots, T-1$ . Find a sequence of policies  $K := \{K_0, \dots, K_{T-1}\}$  to solve:

$$J_{T,p}^*(i_0) := \min_K \mathbb{E}[J_T(K, w, x_0) | i_0]. \quad (5)$$

The policy  $K_t$  at time  $t$  uses information  $i_t$  in (2) and includes  $p \geq 0$  steps of disturbance preview.

We will also consider a similar problem on an infinite horizon. In this case, we will assume the plant  $P$  is linear, time-invariant (LTI):

$$x_{t+1} = A x_t + B_u u_t + B_w w_t, \quad (6)$$

where  $(A, B)$  are a stabilizable pair and  $A$  is nonsingular<sup>1</sup>. We will also assume a sequence of (possibly random) control policies  $K := \{K_0, K_1, \dots\}$  with  $p$  steps of disturbance preview:

$$i_t := \{x_0, \dots, x_t, w_0, \dots, w_{t+p}\}. \quad (7)$$

The cost function for a sequence of policies  $K$  evaluated on a specific disturbance  $w$  and initial condition  $x_0$  is defined as the average per-step cost over the infinite horizon:

$$J_\infty(K, w, x_0) := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} x_t^\top Q x_t + u_t^\top R u_t. \quad (8)$$

Here we assume  $Q \succeq 0$ ,  $R \succ 0$ , and  $(A, Q)$  is detectable. The detectability assumption ensures that any unstable modes in the plant will appear in the cost function. We now state the infinite-horizon (IH), stochastic LQR problem with preview information. We restrict to policies that are stabilizing, i.e. the state remains uniformly bounded in both mean and second-moment.

**Problem 2 (IH Stochastic LQR With Preview).** Consider the LTI plant (6) with  $A$  nonsingular,  $(A, B)$  stabilizable and cost function in (56) with  $Q \succeq 0$ ,  $R \succ 0$  and  $(A, Q)$  detectable. Find a sequence of policies  $K := \{K_0, K_1, \dots\}$  to stabilize the plant and solve:

$$J_{\infty,p}^*(i_0) := \min_{K \text{ stabilizing}} \mathbb{E}[J_\infty(K, w, x_0) | i_0]. \quad (9)$$

The policy  $K_t$  at time  $t$  uses information  $i_t$  in (7) and includes  $p \geq 0$  steps of disturbance preview.

The IH cost could depend on the initial information  $i_0$  for some policies, e.g. a policy could lead to unbounded trajectories and infinite cost for some initial conditions  $x_0$  but not others. However, we will show in Section V-B that the optimal cost does not depend on the initial conditions, i.e.  $J_{\infty,p}^*(i_0) = J_{\infty,p}^*$ .

<sup>1</sup>The solution to the infinite horizon problem (Section V-B) uses a related, discrete-time algebraic Riccati equation (DARE). There are technical issues in solving this DARE when  $A$  is singular. Our infinite horizon results can be extended to the case where  $A$  is singular using generalized solutions for the DARE. See Section 21.3 and Remark 21.2 in [29] for details.

#### IV. PRINCIPLE OF OPTIMALITY

Our solutions for the FH and IH problems with preview will rely on the principle of optimality. This is a well known principle, e.g. see Section 2.2. of [30]. Here we will state a specific version that can be applied to the problems with disturbance preview. We state the principle of optimality for a more general, finite-horizon cost function:

$$J(K, i_0) := \mathbb{E} \left[ g_T(x_T) + \sum_{t=0}^{T-1} g_t(x_t, u_t) \middle| i_0 \right], \quad (10)$$

where the per-step cost functions  $g_0, \dots, g_T$  are given. In this section we drop the subscripts in the cost function  $J$  to simplify the notation. The FH cost in the previous section is a special case where the  $g_0, \dots, g_T$  are quadratic functions. Moreover, we will consider a more general information structure for  $i_t$ . Specifically, we only assume that  $i_t$  is nested, i.e.  $i_t \subseteq i_{t+1}$ . The information structures in the previous section have this nested structure. The goal is to solve  $J^*(i_0) = \min_K J(K, i_0)$ .

We define the value function or optimal cost-to-go at time  $t$  as follows:

$$V_t(i_t) := \min_{K_t, \dots, K_{T-1}} \mathbb{E} \left[ g_T(x_T) + \sum_{j=t}^{T-1} g_j(x_j, u_j) \middle| i_t \right]. \quad (11)$$

The value function at time  $t = 0$  corresponds to the optimal cost:  $V_0(i_0) = J^*(i_0)$ . The principle of optimality states that we can compute the value functions recursively, optimizing over one action at a time.

**Theorem 1** (Principle of Optimality). Define the value function in (11) and assume the information is nested:  $i_0 \subseteq i_1 \subseteq \dots \subseteq i_{T-1}$ . Then the value function satisfies the following backwards recursion starting from  $t = T$ :

$$V_T(i_T) = \mathbb{E}[g_T(x_T) | i_T], \quad (12)$$

$$V_t(i_t) = \min_u \mathbb{E} [g_t(x_t, u_t) + V_{t+1}(i_{t+1}) | i_t, u_t = u] \quad (13)$$

for  $t = 0, \dots, T-1$ .

Moreover, the optimal cost  $J^*(i_0)$  is achieved by a deterministic policy defined by selecting  $u_t = k_t(i_t)$  at each  $t$  to minimize the value function in (13).

*Proof.* Recall the value function defined in (11). Now define a recursive version as

$$W_T(i_T) := \mathbb{E}[g_T(x_T) | i_T],$$

$$W_t(i_t) := \min_u \mathbb{E} [g_t(x_t, u_t) + W_{t+1}(i_{t+1}) | i_t, u_t = u]$$

for  $t = 0, \dots, T-1$ .

We will show, by induction, that  $W_t = V_t$  for all  $t$ . The base case  $W_T = V_T$  holds by definition. Next, make the induction assumption that  $W_{t+1} = V_{t+1}$ . We will prove that  $W_t = V_t$

holds as well. Consider the cost-to-go defined for any fixed set of policies  $\{K_t, \dots, K_{T-1}\}$ :

$$V_t^{K_t: T-1}(i_t) := \mathbb{E} \left[ g_T(x_T) + \sum_{j=t}^{T-1} g_j(x_j, u_j) \middle| i_t \right]. \quad (14)$$

We can re-write this using the tower rule<sup>2</sup>

$$V_t^{K_t: T-1}(i_t) = \mathbb{E} [g_t(x_t, u_t) | i_t] + \mathbb{E} \left[ \mathbb{E} \left[ g_T(x_T) + \sum_{j=t+1}^{T-1} g_j(x_j, u_j) \middle| i_{t+1} \right] \middle| i_t \right]. \quad (15)$$

The tower rule applies due to the assumption of nested information,  $i_t \subseteq i_{t+1}$ . The second term depends on the policies  $\{K_{t+1}, \dots, K_{T-1}\}$ . We can lower bound this term by the value function at time  $t+1$ . This gives:

$$V_t^{K_t: T-1}(i_t) \geq \mathbb{E} [g_t(x_t, u_t) + V_{t+1}(i_{t+1}) | i_t]. \quad (16)$$

We have  $W_{t+1} = V_{t+1}$  by the induction assumption. Next note that the information  $i_t$  is trivially a subset of the information  $\{i_t, u_t\}$ . Hence another application of the tower rule gives:

$$V_t^{K_t: T-1}(i_t) \geq \mathbb{E} \left[ \mathbb{E} [g_t(x_t, u_t) + W_{t+1}(i_{t+1}) | i_t, u_t = u] \middle| i_t \right]. \quad (17)$$

The right side depends on the policy  $K_t$ . As stated previously, this policy is specified by a probability density function  $K_t(u_t, i_t) := \text{Prob}(u_t | i_t)$ . We can replace the outer expectation in (17) by an integral over all control actions:

$$V_t^{K_t: T-1}(i_t) \geq \int \mathbb{E} [g_t(x_t, u_t) + W_{t+1}(i_{t+1}) | i_t, u_t = u] K_t(u, i_t) du. \quad (18)$$

Finally, the right side is lower bounded by minimizing over the control input:

$$V_t^{K_t: T-1}(i_t) \geq \min_u \mathbb{E} [g_t(x_t, u_t) + W_{t+1}(i_{t+1}) | i_t, u_t = u]. \quad (19)$$

The right side is, by definition,  $W_t(i_t)$ . Thus we have shown that  $V_t^{K_t: T-1}(i_t) \geq W_t(i_t)$  for any set of policies  $\{K_t, \dots, K_{T-1}\}$ . This implies that the value function at time  $t$  satisfies  $V_t(i_t) \geq W_t(i_t)$ .

In fact, we can find a set of policies  $\{K_t, \dots, K_{T-1}\}$  that make  $V_t(i_t) = W_t(i_t)$ . For Equation 16, we have equality if we pick  $\{K_{t+1}, \dots, K_{T-1}\}$  to be the optimal policies for the value function at  $t+1$ . For Equation 19, we achieve equality by selecting a deterministic policy  $u_t = k_t(i_t)$  with

$$u_t = \arg \min_u \mathbb{E} [g_t(x_t, u_t) + W_{t+1}(i_{t+1}) | i_t, u_t = u].$$

<sup>2</sup>We have  $\mathbb{E}[\mathbb{E}[x|y]] = \mathbb{E}[x]$ . More generally, whenever  $y \subseteq z$ , we have  $\mathbb{E}[\mathbb{E}[x|z] | y] = \mathbb{E}[x|y]$ . See Theorem 9.1.5 in [31]. This is also known as the law of total expectation.

The policy that achieves  $J^*(i_0)$  from will be deterministic if we use this choice for  $t = 0$  to  $T - 1$ .  $\square$

Note that the proof only relies on the assumption of nested information. It does not use any specific assumptions regarding the model / dynamics relating  $x_{t+1}$  and  $(x_t, u_t)$  (e.g., Markov assumption). It also does not rely on specific structure for the per-step cost functions  $g_0, \dots, g_T$ . That said, we will apply this version of the principle of optimality in the next section for the specific case with linear dynamics, quadratic cost functions, and nested information with disturbance preview.

## V. MAIN RESULTS

### A. FH Stochastic LQR with Preview

This section presents the solution to the FH stochastic LQR problem with disturbance preview (Problem 1). We start with a standard technical lemma regarding a backwards Riccati iteration.

**Lemma 1.** Let  $\{A_t\}_{t=0}^{T-1} \subset \mathbb{R}^{n_x \times n_x}$ ,  $\{B_{u,t}\}_{t=0}^{T-1} \subset \mathbb{R}^{n_x \times n_u}$ ,  $\{Q_t\}_{t=0}^T \subset \mathbb{R}^{n_x \times n_x}$ , and  $\{R_t\}_{t=0}^{T-1} \subset \mathbb{R}^{n_u \times n_u}$  be given. Assume  $Q_t \succeq 0$  and  $R_t \succ 0$  for each  $t$ . Define the following backwards Riccati iteration for  $t = 0, \dots, T$ :

$$\begin{aligned} P_T &:= Q_T, \\ P_t &:= Q_t + A_t^\top P_{t+1} A_t - A_t^\top P_{t+1} B_{u,t} H_t^{-1} B_{u,t}^\top P_{t+1} A_t, \end{aligned} \quad (20)$$

where

$$H_t := R_t + B_{u,t}^\top P_{t+1} B_{u,t} \text{ for } t = 0, \dots, T-1 \quad (21)$$

Then  $P_t \succeq 0$  for  $t = 0, \dots, T$ . and  $H_t \succ 0$ , for  $t = 0, \dots, T-1$ . Hence  $H_t$  is nonsingular for each  $t$  and the Riccati iteration is well defined.

*Proof.* The proof is by induction. The base case is  $P_T \succeq 0$  and  $H_{T-1} \succ 0$ . This follows from the assumptions  $Q_T \succeq 0$  and  $R_{T-1} \succ 0$ . Next, make the induction assumption that  $P_{t+1} \succeq 0$  and  $H_t \succ 0$ . We will show that  $P_t \succeq 0$  and  $H_{t-1} \succ 0$ . Use the generalized matrix inversion lemma (Fact 8.4.13 of [32]) to rewrite the Riccati iteration:

$$P_t = Q_t + A_t^\top P_{t+1} \left( P_{t+1} + P_{t+1} B_{u,t} R_t^{-1} B_{u,t}^\top P_{t+1} \right)^\dagger P_{t+1} A_t,$$

where  $\dagger$  is the Moore-Penrose pseudo inverse. Thus  $P_t \succeq 0$  follows from  $P_{t+1} \succeq 0$ ,  $Q_t \succeq 0$  and  $R_t \succ 0$ . Moreover,  $P_t \succeq 0$  combined with  $R_{t-1} \succ 0$  imply  $H_{t-1} \succ 0$ .  $\square$

The next theorem provides the solution to Problem 1. The proof is based on the principle of optimality (Theorem 1) and dynamic programming.

**Theorem 2.** Consider the FH stochastic LQR with preview including the assumptions stated in Problem 1. Define the following feedback gains using the solution of backwards Riccati iteration (20):

$$K_{x,t} := H_t^{-1} B_{u,t}^\top P_{t+1} A_t, \quad (22)$$

$$K_{w,t} := H_t^{-1} B_{u,t}^\top P_{t+1} B_{w,t}, \quad (23)$$

$$K_{v,t} := H_t^{-1} B_{u,t}^\top. \quad (24)$$

The sequence of policies that achieve  $J_{T,p}^*(i_0)$  are deterministic, and have the form:

$$u_t^* = -K_{x,t} x_t - K_{w,t} w_t - K_{v,t} v_{t+1}, \quad (25)$$

where  $v_{t+1}$  depends on  $\{w_{t+1}, \dots, w_{t+p}\}$  and is given by:

$$\begin{aligned} v_{t+1} &= \sum_{j=t+1}^{t+p} \left[ \hat{A}_{t+1}^\top \cdots \hat{A}_j^\top \right] P_{j+1} B_{w,j} w_j \\ &\text{with } \hat{A}_j := A_j - B_{u,j} K_{x,j}. \end{aligned} \quad (26)$$

Equation 26 uses the convention that  $w_j = 0$  when  $j \geq T$ .

*Proof.* The assumptions in Problem 1 are sufficient for Lemma 1 to hold. Hence the Riccati iteration and the feedback gains (22)-(24) are well-defined. Moreover  $H_t \succ 0$  for each  $t$ .

Define the value function or optimal cost-to-go at time  $t$  for Problem 1 as follows:

$$V_t(i_t) := \min_{K_t, \dots, K_{T-1}} \mathbb{E} \left[ x_T^\top Q_T x_T + \sum_{j=t}^{T-1} x_j^\top Q_j x_j + u_j^\top R_j u_j \mid i_t \right]. \quad (27)$$

The value function at time  $t = 0$  corresponds to the optimal cost:  $V_0(i_0) = J_{T,p}^*(i_0)$ . The value functions can be computed recursively by the principle of optimality (Theorem 1):

$$\begin{aligned} V_T(i_T) &= \mathbb{E}[x_T^\top Q_T x_T \mid i_T] \\ V_t(i_t) &= \min_u \mathbb{E} \left[ x_t^\top Q_t x_t + u_t^\top R_t u_t + V_{t+1}(i_{t+1}) \mid i_t, u_t = u \right] \\ &\text{for } t = 0, \dots, T-1. \end{aligned} \quad (28)$$

We will show that the value functions have the form:

$$V_t(i_t) = x_t^\top P_t x_t + 2\bar{v}_t^\top x_t + q_t \text{ for } t = 0, \dots, T-1, \quad (29)$$

where  $\{P_t\}_{t=0}^T$  satisfy the backwards Riccati recursion (20),  $q_t$  depend on  $\{w_t, \dots, w_{t+p}\}$ , and  $\bar{v}_t$  is given by:

$$\bar{v}_t := \sum_{j=t}^{t+p} \left[ \hat{A}_t^\top \cdots \hat{A}_j^\top \right] P_{j+1} B_{w,j} w_j. \quad (30)$$

Equation 30 again uses the convention that  $w_j = 0$  if  $j \geq T$ .<sup>3</sup> We will show that (29) holds by induction. The base case at  $t = T$  corresponds to  $P_T = Q_T$ ,  $\bar{v}_T = 0$ , and  $q_T = 0$ . Next make the induction assumption that the value function at time  $t+1$  has the given form:

$$V_{t+1}(i_{t+1}) = x_{t+1}^\top P_{t+1} x_{t+1} + 2\bar{v}_{t+1}^\top x_{t+1} + q_{t+1}. \quad (31)$$

We will prove that the value function at time  $t$  has the similar form given in (29). Substitute (31) into the value function recursion (28) and use the LTV dynamics (1) to replace  $x_{t+1}$ . This yields:

<sup>3</sup>Note that  $\bar{v}_t$  depends on  $\{w_t, \dots, w_{t+p}\}$ . This means that  $\bar{v}_{t+1}$  depends on  $\{w_{t+1}, \dots, w_{t+p+1}\}$ . However,  $v_{t+1}$  defined in (26) depends on  $\{w_{t+1}, \dots, w_{t+p}\}$ . Hence  $\bar{v}_{t+1}$  and  $v_{t+1}$  differ by one term.

$$\begin{aligned}
V_t(i_t) &= \min_u \mathbb{E}[x_t^\top Q_t x_t + u_t^\top R_t u_t] \\
&+ (A_t x_t + B_{u,t} u_t + B_{w,t} w_t)^\top P_{k+1} (A_t x_t + B_{u,t} u_t + B_{w,t} w_t) \\
&+ 2\bar{v}_{t+1}^\top (A_t x_t + B_{u,t} u_t + B_{w,t} w_t) + q_{t+1} | i_t, u_t = u].
\end{aligned} \tag{32}$$

Note that the quadratic term in this minimization is  $u_t^\top H_t u_t$  where  $H_t \succ 0$ . Thus the minimization has a unique minima because the objective is a strictly convex function. Take the gradient respect to  $u$  to find the optimal input:

$$u_t^* = \mathbb{E}[-K_{x,t} x_t - K_{w,t} w_t - K_{v,t} \bar{v}_{t+1} | i_t]. \tag{33}$$

where the gains are defined in (22)-(24). Note that  $(x_t, w_t)$  are contained in  $i_t$  so  $E[x_t | i_t] = x_t$  and  $E[w_t | i_t] = w_t$ . Moreover, the induction assumption gives:

$$\bar{v}_{t+1} = \sum_{j=t+1}^{t+p+1} \left[ \hat{A}_{t+1}^\top \cdots \hat{A}_j^\top \right] P_{j+1} B_{w,j} w_j. \tag{34}$$

Thus  $E[\bar{v}_{t+1} | i_t] = v_{t+1}$  and the optimal input in (33) simplifies to the expression in (25).

Next, substitute  $u_t^*$  back in to the cost-to-go function (32). After some simplification, the value function at time  $t$  has the form given in (29) with  $P_t$  given by the backwards Riccati recursion (20) and  $\bar{v}_t$  given by (30). Moreover,  $q_t$  depends on  $\{w_t, \dots, w_{t+p}\}$  and can be simplified to:

$$\begin{aligned}
q_t &= -(K_{w,t} w_t + K_{v,t} v_{t+1})^\top H_t (K_{w,t} w_t + K_{v,t} v_{t+1}) \\
&+ w_t^\top B_{w,t}^\top (P_{t+1} B_{w,t} w_t + 2v_{t+1}) + \mathbb{E}[q_{t+1} | i_t].
\end{aligned}$$

Hence the proof is complete by induction.  $\square$

The optimal cost, based on the proof, is

$$J_{T,p}^*(i_0) = x_0^\top P_0 x_0 + 2\bar{v}_0^\top x_0 + q_0. \tag{35}$$

The first term depends only on the state initial condition  $x_0$ . The second term depends on both the  $x_0$  and the initial disturbance information  $\{w_0, \dots, w_p\}$ . The third term also depends on the initial disturbance information but also includes the expected cost of the disturbances  $\{w_{p+1}, \dots, w_{T-1}\}$ .

### B. IH Stochastic LQR with Preview

This section presents the solution to the IH stochastic LQR problem with disturbance preview (Problem 2). We start with a standard technical lemma regarding a discrete time algebraic Riccati equation (DARE).

**Lemma 2.** Let  $A \in \mathbb{R}^{n_x \times n_x}$ ,  $B_u \in \mathbb{R}^{n_x \times n_u}$ ,  $Q \in \mathbb{R}^{n_x \times n_x}$ , and  $R \in \mathbb{R}^{n_u \times n_u}$  be given. Assume: (i)  $Q \succeq 0$  and  $R \succ 0$ , (ii)  $(A, B_u)$  stabilizable, (iii)  $A$  is nonsingular, and (iv)  $(A, Q)$  has no unobservable modes on the unit circle. Then there is a unique stabilizing solution  $P \succeq 0$  such that:

1)  $P$  satisfies the following DARE:

$$0 = P - A^\top P A - Q + A^\top P B_u H^{-1} B_u^\top P A, \tag{36}$$

where  $H := R + B_u^\top P B_u \succ 0$ .

2) The gain  $K_x := H^{-1} B_u^\top P A$  is stabilizing, i.e.  $A - B_u K_x$  is a Schur matrix.

*Proof.* Statements 1) and 2) follow from Corollary 21.13 and Theorem 21.7 of [29] (after aligning the notation).  $\square$

The next theorem constructs the IH stochastic LQR with preview controller using the stabilizing solution of the DARE.

**Theorem 3.** Consider the IH stochastic LQR with preview including the assumptions stated in Problem 2. Define the following feedback gains using the solution of DARE (36):

$$K_x := H^{-1} B_u^\top P A, \tag{37}$$

$$K_w := H^{-1} B_u^\top P B_w, \tag{38}$$

$$K_v := H^{-1} B_u^\top. \tag{39}$$

The sequence of policies that achieve  $J_{\infty,p}^*(i_0)$  are deterministic, and have the form:

$$u_t^* = -K_x x_t - K_w w_t - K_v v_{t+1}, \tag{40}$$

where  $v_{t+1}$  depends on  $\{w_{t+1}, \dots, w_{t+p}\}$  and is given by:

$$\begin{aligned}
v_{t+1} &= \sum_{j=t+1}^{t+p} \left[ \hat{A}^\top \right]^{j-t} P B_w w_j \\
&\text{with } \hat{A} := A - B_u K_x.
\end{aligned} \tag{41}$$

Equation 41 uses the convention that  $w_j = 0$  when  $j \geq T$ .

*Proof.* Recall that the infinite horizon cost for any sequence of policies  $K := \{K_0, K_1, \dots\}$  is given by

$$J_\infty(K, w, x_0) := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} g(x_t, u_t), \tag{42}$$

where  $g(x_t, u_t) := x_t^\top Q x_t + u_t^\top R u_t$  is the per-step cost. The main part of the proof is to re-arrange the per-step cost into a form that involves  $u_t - u_t^*$ . This will be used to show that  $u_t = u_t^*$  is the optimal input.

First, substitute for  $Q$  using the DARE (36) and re-arrange to express the per-step cost as follows:

$$\begin{aligned}
g(x_t, u_t) &= (u_t + K_x x_t)^\top H (u_t + K_x x_t) + x_t^\top P x_t \\
&- (A x_t + B_u u_t)^\top P (A x_t + B_u u_t).
\end{aligned}$$

Use the LTI dynamics (6) to replace  $A x_t + B_u u_t$  by  $x_{t+1} - B_w w_t$ . This yields

$$\begin{aligned}
g(x_t, u_t) &= (u_t + K_x x_t)^\top H (u_t + K_x x_t) \\
&+ (x_t^\top P x_t - x_{t+1}^\top P x_{t+1}) - w_t^\top B_w^\top P B_w w_t + 2x_{t+1}^\top P B_w w_t.
\end{aligned}$$

Next, use the definition of  $u_t^*$  in (40) to replace  $K_x x_t$  with  $-u_t^* - K_w w_t - K_v v_{t+1}$ . The per-step cost is thus given by

$$\begin{aligned}
g(x_t, u_t) &= (u_t - u_t^*)^\top H (u_t - u_t^*) + (x_t^\top P x_t - x_{t+1}^\top P x_{t+1}) + l_t \\
&- w_t^\top B_w^\top P B_w w_t - (K_w w_t + K_v v_{t+1})^\top H (K_w w_t + K_v v_{t+1}),
\end{aligned} \tag{43}$$

where

$$l_t := -2(K_w w_t + K_v v_{t+1})^\top H (u_t + K_x x_t) + 2x_{t+1}^\top P B_w w_t.$$

We now focus on simplifying the term  $l_t$ . We have, from (38) and (39), that  $HK_w = B_u^\top PB_w$  and  $HK_v = B_u^\top$ . Hence  $l_t$  simplifies to:

$$l_t := -2(PB_w w_t + v_{t+1})^\top B_u (u_t + K_x x_t) + 2x_{t+1}^\top PB_w w_t.$$

We can replace  $B_u (u_t + K_x x_t)$  with  $x_{t+1} - \hat{A}x_t - B_w w_t$ . Combining terms yields:

$$l_t := 2(PB_w w_t + v_{t+1})^\top (\hat{A}x_t + B_w w_t) - 2x_{t+1}^\top v_{t+1}.$$

Note that the preview term  $v_{t+1}$  defined in (41) satisfies the following relation:

$$\hat{A}^\top (PB_w w_t + v_{t+1}) = v_t + [\hat{A}^\top]^{p+1} PB_w w_{t+p}. \quad (44)$$

We can use this to finally simplify  $l_t$  to:

$$l_t = 2x_t^\top v_t + 2x_t^\top [\hat{A}^\top]^{p+1} PB_w w_{t+p} + 2w_t^\top B_w^\top PB_w w_t + 2v_{t+1}^\top B_w w_t - 2x_{t+1}^\top v_{t+1}.$$

Substitute this expression for  $l_t$  back into the per-step cost (43). Re-organizing terms gives:

$$g(x_t, u_t) = (u_t - u_t^*)^\top H (u_t - u_t^*) \quad (45a)$$

$$- (K_w w_t + K_v v_{t+1})^\top H (K_w w_t + K_v v_{t+1}) \quad (45b)$$

$$+ (x_t^\top P x_t - x_{t+1}^\top P x_{t+1}) + 2(x_t^\top v_t - x_{t+1}^\top v_{t+1}) \quad (45c)$$

$$+ 2x_t^\top [\hat{A}^\top]^{p+1} PB_w w_{t+k} \quad (45d)$$

$$+ 2v_{t+1}^\top B_w w_t. \quad (45e)$$

The terms (45c) form telescoping sums when inserted into the cost  $J_\infty(K, w, x_0)$ . The telescoping sums contribute zero to the cost as  $T \rightarrow \infty$  since the policy is assumed to be stabilizing. The expectation of term (45e) is zero as  $v_{t+1}$  depends on  $\{w_{t+1}, \dots, w_{t+p}\}$  and is independent of  $w_t$ . The expectation of term (45d) is also zero as  $x_t$  depends on information  $i_{t-1} = \{x_0, \dots, x_{t-1}, w_0, \dots, w_{t+p-1}\}$ , which is independent of  $w_{t+p}$ . Thus the cost for any stabilizing policy simplifies to:

$$J_{\infty,p}^*(i_0) = \min_{K \text{ stabilizing}} \mathbb{E} \left[ \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} (u_t - u_t^*)^\top H (u_t - u_t^*) - (K_w w_t + K_v v_{t+1})^\top H (K_w w_t + K_v v_{t+1}) + w_t^\top B_w^\top PB_w w_t \middle| i_0 \right]. \quad (46)$$

Only the first term depends on the plant input and  $H \succ 0$ . Thus  $J_{\infty,p}^*(i_0)$  is achieved by  $u_t = u_t^*$ .  $\square$

The optimal IH cost is obtained by substituting  $u_t = u_t^*$  into (46). This yields:

$$J_{\infty,p}^*(i_0) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=0}^{T-1} w_t^\top B_w^\top PB_w w_t - (K_w w_t + K_v v_{t+1})^\top H (K_w w_t + K_v v_{t+1}) \middle| i_0 \right].$$

This can be simplified further using the definition of  $v_{t+1}$  in (41) combined with  $\mathbb{E}[w_i w_j^\top] = \delta_{ij} I$ . This gives:

$$J_{\infty,p}^* = \text{trace} [PB_w B_w^\top] - \text{trace} \left[ \sum_{j=0}^p HK_v (\hat{A}^\top)^j PB_w B_w^\top P \hat{A}^j K_v^\top \right]. \quad (47)$$

Here we have used  $K_w = K_v PB_w$  to simplify the expression further. Note that the optimal IH cost does not depend on the initial condition, i.e.  $J_{\infty,p}^*(i_0) = J_{\infty,p}^*$ . The first term in the optimal cost,  $\text{trace}[PB_w B_w^\top]$ , is the cost that would be achieved using only state feedback. The second term gives the cost reduction obtained by using the current disturbance measurement and  $p$  steps of preview.

### C. Relation to Optimal Noncausal Controller

This section discusses the optimal noncausal controller with full knowledge of the (past, current and future) values of the disturbance. We focus on the IH case where the controller has access to the following information at time  $t$ :

$$i_t := \{x_0, \dots, x_t, w_0, w_1, \dots\}. \quad (48)$$

The disturbance is fully known so we consider deterministic policies  $K := \{k_0, k_1, \dots\}$  where  $u_t = k_t(i_t)$ . Moreover, we assume  $w \in \ell_2$  and the (deterministic) cost is:

$$\sum_{t=0}^{\infty} x_t^\top Q x_t + u_t^\top R u_t. \quad (49)$$

Here the cost is deterministic (in contrast to the expected costs defined previously). The next problem states the IH design with full preview.

**Problem 3** (IH LQR With Full Preview). Consider the LTI plant (6) with  $A$  nonsingular,  $(A, B)$  stabilizable and cost function in (49) with  $Q \succeq 0$ ,  $R \succ 0$  and  $(A, Q)$  detectable. Also assume  $w \in \ell_2$ . Find a sequence of deterministic policies  $K := \{k_0, k_1, \dots\}$  to stabilize the plant and solve:

$$\min_{K \text{ stabilizing}} \sum_{t=0}^{\infty} x_t^\top Q x_t + u_t^\top R u_t. \quad (50)$$

The policy  $k_t$  at time  $t$  uses information  $i_t$  in (48) and includes full preview of the disturbance.

A solution for the optimal noncausal controller is given in Theorem 11.2.1 of [28]. Related noncausal results (both FH and IH) are given in [1], [2], [3], [4], [5] where the noncausal controller is used as a baseline for regret-based control design. The next result gives the IH noncausal controller in a form/notation that closely aligns with the finite preview controller in Theorem 3.

**Theorem 4.** Consider IH LQR with full preview including the assumptions stated in Problem 3. Define the feedback gains  $(K_x, K_w, K_v)$  in (37)-(39) using the solution of DARE

(36). The sequence of policies that achieve the minimum are deterministic and have the form:

$$u_t^{nc} = -K_x x_t - K_w w_t - K_v v_{t+1}^{nc}, \quad (51)$$

where  $v_{t+1}^{nc}$  depends on  $\{w_{t+1}, w_{t+2}, \dots\}$  and is given by the following anticausal dynamics:

$$\begin{aligned} v_t^{nc} &= \hat{A}^\top (v_{t+1}^{nc} + PB_w w_t), \quad v_\infty^{nc} = 0 \\ \text{with } \hat{A} &:= A - B_u K_x. \end{aligned} \quad (52)$$

This result, as stated, is a special case of Theorem 1 in [33]. It was stated in [33] for a more general LQR cost including a cross term  $2x_t^\top S w_t$ .<sup>4</sup> Note that the full preview design (Problem 3) was stated with a deterministic formulation while finite preview design (Problem 2) was derived with stochastic formulation. However, we show next that the controller with finite preview  $p < \infty$  converges, in the limit as  $p \rightarrow \infty$ , to the optimal noncausal controller for any fixed disturbance  $w \in \ell_2$ .

**Theorem 5.** Let  $w \in \ell_2$  be any given (deterministic) disturbance sequence. Then the optimal noncausal controller is given by  $u_t^{nc}$  in (51) where  $v_{t+1}^{nc}$  can be expressed as:

$$v_{t+1}^{nc} = \sum_{j=t+1}^{\infty} \left[ \hat{A}^\top \right]^{j-t} PB_w w_j \quad (53)$$

Moreover, let  $u_t^p$  denote the optimal finite preview controller given in (40) with  $p < \infty$  steps of preview. Then  $\|u_t^{nc} - u_t^p\|_2 \rightarrow 0$  uniformly in  $t$  as  $p \rightarrow \infty$ .

*Proof.* We will first verify that (53) holds. Specifically, we will show that  $v_{t+1}^{nc}$  defined by (52) is equal to the following signal:

$$y_{t+1} = \sum_{j=t+1}^{\infty} \left[ \hat{A}^\top \right]^{j-t} PB_w w_j \quad (54)$$

It can be verified, by direct substitution, that  $y_{t+1}$  satisfies the state-space recursion in (52). Moreover,  $y_{t+1}$  satisfies the boundary condition  $y_\infty = 0$  because  $w \in \ell_2$  and  $\hat{A}$  is a Schur matrix.<sup>5</sup> Thus  $y_{t+1}$  is the unique solution to the noncausal dynamics in (52) and hence (53) holds.

Next, we show that  $\|u_t^{nc} - u_t^p\|_2 \rightarrow 0$  uniformly in  $t$  as  $p \rightarrow \infty$ . The difference between the two commands is:

$$u_t^{nc} - u_t^p = -K_v \cdot \sum_{j=t+p+1}^{\infty} \left[ \hat{A}^\top \right]^{j-t} PB_w w_j \quad (55)$$

We can bound this difference as  $\|u_t^{nc} - u_t^p\|_2 \leq a_p \cdot b$  where

$$a_p := \left\| \sum_{l=p+1}^{\infty} \left[ \hat{A}^\top \right]^l PB_w \right\|_{2 \rightarrow 2} \quad \text{and } b := \max_j \|w_j\|_2$$

<sup>4</sup>The proof given in [33] was for signals defined for  $t$  from  $-\infty$  to  $+\infty$ . However, the optimal noncausal controller is the same for signals defined from  $t = 0$  to  $\infty$ .

<sup>5</sup>We can bound  $y_{t+1}$  by  $\|y_{t+1}\|_2 \leq a \cdot b_t$  where  $a := \left\| \sum_{l=1}^{\infty} \left[ \hat{A}^\top \right]^l PB_w \right\|_{2 \rightarrow 2}$  and  $b_t := \max_{j \geq t+1} \|w_j\|_2$ . We have  $a < \infty$  because  $\hat{A}$  is a Schur matrix. Moreover,  $b_t \rightarrow 0$  as  $t \rightarrow \infty$  because  $w \in \ell_2$ . Hence  $y_{t+1} \rightarrow 0$  as  $t \rightarrow \infty$ .

The assumption  $w \in \ell_2$  implies that  $b$  is finite. Moreover,  $\hat{A}$  is a Schur matrix and hence  $a_p \rightarrow 0$  as  $p \rightarrow \infty$ . This implies that  $\|u_t^{nc} - u_t^p\|_2 \rightarrow 0$  uniformly in  $t$  as  $p \rightarrow \infty$ .  $\square$

Finally, we can derive a related convergence result for the costs achieved with finite preview and full, noncausal preview. The optimal noncausal controller depends on the specific (deterministic) disturbance sequence and assumes  $w \in \ell_2$  to ensure the cost (49) is bounded. However, we can apply the noncausal controller to stochastic disturbances  $w$  that are drawn from the distribution of IID, zero mean, unit variance sequences. The expected per-step cost for the optimal noncausal controller is:

$$J_{\infty,nc}^*(x_0) := \mathbb{E}[J_\infty(K_{nc}, w, x_0) | x_0], \quad (56)$$

where  $K_{nc}$  is the optimal noncausal controller (51) for a specific sequence  $w$ . The expected per-step cost for the noncausal controller is averaged over stochastic signals  $w$ . We show below that it does not depend on the initial state, i.e.  $J_{\infty,nc}^*(x_0) = J_{\infty,nc}^*$ . The next result is that the expected per-step cost with the optimal finite preview controller converges to the cost achieved by the optimal noncausal controller. We also give a simple expression to compute the optimal cost achieved by the noncausal controller.

**Theorem 6.** Let  $w$  be an IID sequence with  $\mathbb{E}[w_i] = 0$  and  $\mathbb{E}[w_i w_j^\top] = \delta_{ij} I$  for  $i, j \in \{0, 1, \dots\}$ . The expected per-step cost for the optimal noncausal controller is:

$$J_{\infty,nc}^* = \text{trace}[PB_w B_w^\top] - \text{trace}[HK_v X K_v^\top]. \quad (57)$$

where  $X$  is the solution to the following discrete-time Lyapunov equation:

$$\hat{A}^\top X \hat{A} - X + PB_w B_w^\top P = 0 \quad (58)$$

Moreover, let  $J_{\infty,p}^*$  denote the optimal expected per-step cost in (47) achieved by the optimal IH controller with  $p < \infty$  steps of preview. Then  $J_{\infty,p}^* \rightarrow J_{\infty,nc}^*$  as  $p \rightarrow \infty$ .

*Proof.* The steps in the proof of Theorem 3 can be used to show that the noncausal cost is given by:

$$\begin{aligned} J_{\infty,nc}^* &= \text{trace}[PB_w B_w^\top] \\ &\quad - \text{trace} \left[ \sum_{j=0}^{\infty} HK_v (\hat{A}^\top)^j PB_w B_w^\top P \hat{A}^j K_v^\top \right]. \end{aligned} \quad (59)$$

The infinite sum converges because  $\hat{A}$  is a Schur matrix by Lemma 2. Moreover, the difference between the optimal finite preview and noncausal costs is:

$$J_{\infty,p}^* - J_{\infty,nc}^* = \text{trace} \left[ \sum_{j=p+1}^{\infty} HK_v (\hat{A}^\top)^j PB_w B_w^\top P \hat{A}^j K_v^\top \right]. \quad (60)$$

The convergence  $J_{\infty,p}^* \rightarrow J_{\infty,nc}^*$  follows because  $\hat{A}$  is a Schur matrix. Finally, define  $X := \sum_{j=0}^{\infty} (\hat{A}^\top)^j PB_w B_w^\top P \hat{A}^j$ .  $X$  is

the solution of the discrete-time Lyapunov equation (58) (See Section 21.1 in [29]). This yields the expression for the optimal cost in (57).  $\square$

A consequence of the previous results is that the optimal finite preview cost converges geometrically to the optimal noncausal cost as  $p \rightarrow \infty$ . This is formally stated next.

**Corollary 1.** Let  $w$  be an IID sequence with  $\mathbb{E}[w_i] = 0$  and  $\mathbb{E}[w_i w_j^\top] = \delta_{ij} I$  for  $i, j \in \{0, 1, \dots\}$ . The difference between the optimal finite preview and noncausal costs is:

$$J_{\infty,p}^* - J_{\infty,nc}^* = \text{trace} \left[ Y (\hat{A}^\top)^{p+1} X (\hat{A})^{p+1} \right]. \quad (61)$$

where  $Y := K_v^\top H K_v \succeq 0$ . This difference is bounded by:

$$\begin{aligned} \lambda_{\min}(X^{\frac{1}{2}})^2 \lambda_{\min}(Y^{\frac{1}{2}})^2 \|\hat{A}^{p+1}\|_F^2 &\leq J_{\infty,p}^* - J_{\infty,nc}^* \\ &\leq \|X^{\frac{1}{2}}\|_F^2 \|Y^{\frac{1}{2}}\|_F^2 \|\hat{A}^{p+1}\|_F^2. \end{aligned} \quad (62)$$

*Proof.* The difference between the optimal finite preview and noncausal costs is given in Equation 60. This expression can be rewritten as follows:

$$\begin{aligned} J_{\infty,p}^* - J_{\infty,nc}^* = \\ \text{trace} \left[ Y (\hat{A}^\top)^{p+1} \left( \sum_{j=0}^{\infty} (\hat{A}^\top)^j P B_w B_w^\top P \hat{A}^j \right) \hat{A}^{p+1} \right]. \end{aligned}$$

The summation in parentheses is, by definition, equal to  $X$  and this yields Equation 61. Next, rewrite this expression in terms of the Frobenius norm: (61):

$$J_{\infty,p}^* - J_{\infty,nc}^* = \|X^{1/2} \hat{A}^{p+1} Y^{1/2}\|_F^2.$$

We obtain the upper bound in (62) using the submultiplicative property of the Frobenius norm:  $\|PQ\|_F \leq \|P\|_F \|Q\|_F$  for any two matrices  $P$  and  $Q$  of appropriate dimensions. We obtain the lower bound in (62) using the following property:  $\|PQ\|_F \geq \lambda_{\min}(P) \|Q\|_F$  where  $P \succeq 0$  and  $Q$  is any matrix of appropriate dimension.  $\square$

The spectral radius formula gives the asymptotic approximation  $\lim_{k \rightarrow \infty} \|\hat{A}^k\|_F^{\frac{1}{k}} = \rho(\hat{A})$ . Thus the asymptotic bounds in (62) as  $p \rightarrow \infty$  are approximated by:

$$c_1 \rho(\hat{A})^{2(p+1)} \lesssim J_{\infty,p}^* - J_{\infty,nc}^* \lesssim c_2 \rho(\hat{A})^{2(p+1)} \quad (63)$$

where  $c_1 := \lambda_{\min}(X) \lambda_{\min}(Y)$  and  $c_2 := \text{trace}(X) \text{trace}(Y)$  are constants independent of  $p$ . This implies that the finite preview cost converges to the noncausal cost geometrically with  $p$ .

## VI. EXAMPLE

This section presents a simple example to illustrate the solutions to the FH and IH stochastic LQR problems. Consider a simplified LTI model for the longitudinal dynamics of a Boeing 747 linearized at one steady, level flight condition (Problem 17.6 in [34]):

$$x_{t+1} = A x_t + B_u u_t + w_t \quad (64)$$

where the state matrices are:

$$A = \begin{bmatrix} 0.99 & 0.03 & -0.02 & -0.32 \\ 0.01 & 0.47 & 4.7 & 0 \\ 0.02 & -0.06 & 0.4 & 0 \\ 0.01 & -0.04 & 0.72 & 0.99 \end{bmatrix} \quad \text{and} \quad B_u = \begin{bmatrix} 0.01 & 0.99 \\ -3.44 & 1.66 \\ -0.83 & 0.44 \\ -0.47 & 0.25 \end{bmatrix}.$$

We consider the IH and FH stochastic LQR problems with constant cost matrices  $Q = I$  and  $R = I$  at each step. This example was used in prior work on competitive ratio [6] and additive regret [1], [2].

Figure 1 shows the performance achieved by the optimal FH controller on the horizon  $T = 100$  for previews  $p = 5$  and 20 (solid lines). A closed-loop simulation was performed with a random disturbance. The plot shows the per-step cost averaged from time 0 to time  $t$  versus the time  $t$ . This is compared against the performance of the optimal  $H_2$  controller computed on the IH (dashed lines). This  $H_2$  controller is computed using a system augmented with delays to account for the disturbance preview as in [19], [7], [25], [23], [24], [21]. This full information  $H_2$  controller is computed using `h2syn` in Matlab. The optimal FH controller and IH  $H_2$  controller have similar performance for  $t \leq 60$ . In fact, the FH and IH gains are similar at the beginning of the simulation. However, the FH gains deviate significantly from the IH gains near the end of the horizon. The FH gains are optimal for the horizon and hence the FH controller yields improved performance (lower average cost) by the end of the horizon.

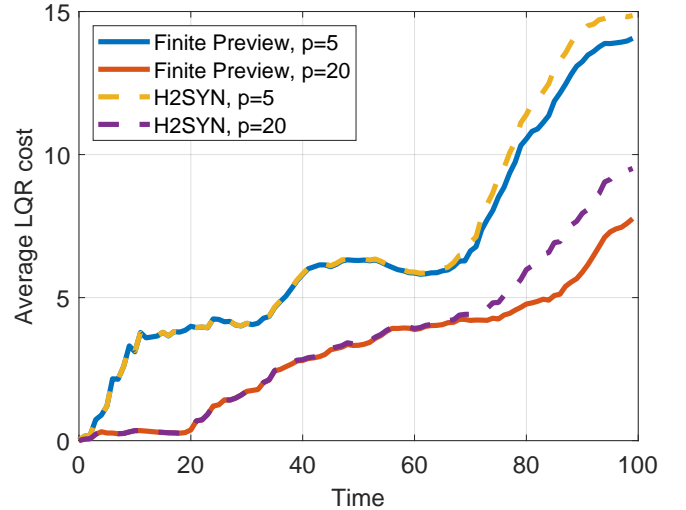


Figure 1. Comparison of average cost obtained by optimal FH controller and optimal  $H_2$  controller computed via an augmented system.

Figure 2 shows the optimal IH cost versus the preview horizon  $p$  (blue solid). The cost achieved by the finite preview controller is computed using the expression in Equation 47. The plot also shows the cost achieved by the Linear Quadratic Regulator (LQR) state feedback:  $u_t = -K_x x_t$  where  $K_x = H^{-1} B_u^\top P A$  is the same gain as used in the finite preview controller. The LQR state feedback achieves a cost equal to  $\text{trace}[P B_w B_w^\top] \approx 33.2$  (red circle). The full information controller  $u_t = -K_x x_t - K_w w_t$  corresponds to the case of  $p = 0$ . This improves on the LQR performance and reduces the cost to  $\approx 29.4$ . Adding disturbance preview ( $p > 1$ ) further



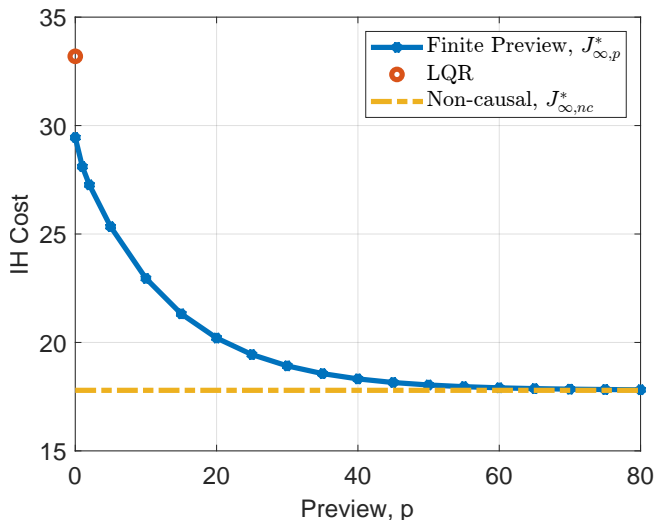


Figure 2. Comparison of IH cost versus preview horizon  $p$ .

decreases the cost. In fact, the cost monotonically decreases with increasing preview.

Figure 2 also shows the performance of the optimal non-causal controller (Theorem 4) for comparison. The cost for the optimal non-causal controller depends on the specific disturbance input. The yellow dashed line in Figure 2 corresponds to the expected (average) cost of the non-causal controller assuming white noise disturbances. This was computed using the expression given in Theorem 6. The noncausal controller has full knowledge of the disturbance and yields the lowest possible cost of  $\approx 17.8$ . The performance of the finite preview controller converges, as  $p \rightarrow \infty$ , to the performance of the optimal non-causal controller. This is expected based on the discussion in Section V-C.

Finally, Figure 3 shows the optimal IH relative error of the cost for the finite preview controller compared to the cost for the optimal noncausal controller. This is normalized by the noncausal cost and plotted on a logarithmic scale against the preview horizon  $p$  (blue solid). The plot uses the same cost results as in Figure 2. The plot is linear on a log scale in agreement with the geometric convergence given in Corollary 1. The slope on the log scale should be approximately  $2 \log \rho(A)$  as  $p \rightarrow \infty$ .

## VII. CONCLUSION

This paper presented solutions to the discrete-time, stochastic LQR problem with  $p$  steps of disturbance preview information where  $p$  is finite. Solutions were presented for both the finite horizon and infinite horizon problems. Moreover, we provided a proof for the principle of optimality that relies only on the assumption of nested information structure. This is of independent interest. Finally, we compared our solutions to several existing results for optimal control with preview information. Future work will explore using our finite-preview controller as a baseline for regret-based design.

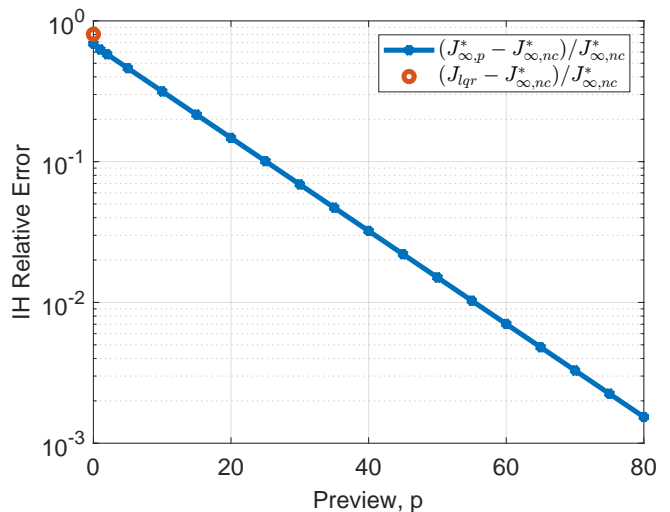


Figure 3. Comparison of IH relative error on a log scale versus preview horizon  $p$ .

## VIII. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 2347026. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- [1] O. Sabag, G. Goel, S. Lale, and B. Hassibi, "Regret-optimal controller for the full-information problem," in *American Control Conference*, 2021, pp. 4777–4782.
- [2] —, "Regret-optimal full-information control," *arXiv preprint arXiv:2105.01244*, 2021.
- [3] O. Sabag, S. Lale, and B. Hassibi, "Optimal competitive-ratio control," *arXiv preprint arXiv:2206.01782*, 2022.
- [4] G. Goel and B. Hassibi, "Regret-optimal control in dynamic environments," *arXiv preprint arXiv:2010.10473*, 2020.
- [5] —, "The power of linear controllers in LQR control," in *IEEE Conference on Decision and Control*. IEEE, 2022, pp. 6652–6657.
- [6] —, "Competitive control," *IEEE Transactions on Automatic Control*, vol. 68, no. 9, pp. 5162–5173, 2022.
- [7] A. A. Moelja and G. Meinsma, " $H_2$  control of preview systems," *Automatica*, vol. 42, no. 6, pp. 945–952, 2006.
- [8] G. Marro and E. Zattoni, " $H_2$ -optimal rejection with preview in the continuous-time domain," *Automatica*, vol. 41, no. 5, pp. 815–821, 2005.
- [9] C. Sentouh, B. Soualmi, J. C. Popieul, and S. Debernard, "The  $H_2$ -optimal preview controller for a shared lateral control," in *International Conference on Intelligent Transportation Systems*, 2011, pp. 1452–1458.
- [10] M. Tomizuka, "Optimal continuous finite preview problem," *IEEE Transactions on Automatic Control*, vol. 20, no. 3, pp. 362–365, 1975.
- [11] A. Kojima and S. Ishijima, "LQ preview synthesis: optimal control and worst case analysis," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 352–357, 1999.
- [12] A. Lindquist, "On optimal stochastic control with smoothed information," *Information Sciences*, vol. 1, no. 1, pp. 55–85, 1968.
- [13] A. Hać, "Optimal linear preview control of active vehicle suspension," *Vehicle System Dynamics*, vol. 21, no. 1, pp. 167–195, 1992.
- [14] H. Peng and M. Tomizuka, "Optimal preview control for vehicle lateral guidance," California PATH, Tech. Rep. UCB-ITS-PRR-91-16, 1991.
- [15] T. Yoshimura, K. Edokoro, and N. Ananthanarayana, "An active suspension model for rail/vehicle systems with preview and stochastic optimal control," *J. of Sound and Vibration*, vol. 166, no. 3, pp. 507–519, 1993.
- [16] J. Marzbanrad, G. Ahmadi, H. Zohoor, and Y. Hojjat, "Stochastic optimal preview control of a vehicle suspension," *Journal of Sound and Vibration*, vol. 275, no. 3, pp. 973–990, 2004.

- [17] K. Hashikura, R. Hotchi, A. Kojima, and T. Masuta, "On implementations of  $H_2$  preview output feedback law with application to LFC with load demand prediction," *International Journal of Control*, vol. 93, no. 4, pp. 844–857, 2020.
- [18] A. J. Hazell and D. J. N. Limebeer, "A framework for discrete-time  $H_2$  preview control," *Journal of Dynamic Systems, Measurement, and Control*, vol. 132, no. 3, p. 031005, 04 2010.
- [19] K. Ting, M. Mesbahi, E. Livne, and K. A. Morgansen, "Wind tunnel study of preview  $H_2$  and  $H_\infty$  control for gust load alleviation for flexible aircraft," *AIAA SciTech 2022 Forum*, 2022.
- [20] E. Zattoni, " $H_2$ -optimal decoupling with preview: a dynamic feedforward solution based on factorization techniques," in *American Control Conference*, 2006, pp. 316–320.
- [21] M. Tomizuka and D. E. Whitney, "Optimal Discrete Finite Preview Problems (Why and How Is Future Information Important?)," *Journal of Dynamic Systems, Measurement, and Control*, vol. 97, no. 4, pp. 319–325, 12 1975.
- [22] M. Tomizuka, "Optimum Linear Preview Control With Application to Vehicle Suspension"—Revisited," *Journal of Dynamic Systems, Measurement, and Control*, vol. 98, no. 3, pp. 309–315, 09 1976.
- [23] H.-S. Roh and Y. Park, "Stochastic optimal preview control of an active vehicle suspension," *J. of Sound and Vibration*, vol. 220, no. 2, pp. 313–330, 1999.
- [24] N. Louam, D. A. Wilson, and R. S. Sharp, "Optimal control of a vehicle suspension incorporating the time delay between front and rear wheel inputs," *Vehicle System Dynamics*, vol. 17, no. 6, pp. 317–336, 1988.
- [25] S. Yim, "Design of a preview controller for vehicle rollover prevention," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4217–4226, 2011.
- [26] M. Tomizuka and D. E. Rosenthal, "On the Optimal Digital State Vector Feedback Controller With Integral and Preview Actions," *Journal of Dynamic Systems, Measurement, and Control*, vol. 101, no. 2, pp. 172–178, 06 1979.
- [27] M. Tomizuka and D. H. Fung, "Design of Digital Feedforward/Preview Controllers for Processes With Predetermined Feedback Controllers," *Journal of Dynamic Systems, Measurement, and Control*, vol. 102, no. 4, pp. 218–225, 12 1980.
- [28] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic estimation and control: a unified approach to  $H_2$  and  $H_\infty$  theories*. SIAM, 1999.
- [29] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control*. Prentice hall, 1996.
- [30] B. Anderson and J. Moore, *Optimal control: linear quadratic methods*. Prentice-Hall, 1990.
- [31] K. Chung, *A Course in Probability Theory*. Academic Press, 2001.
- [32] D. Bernstein, *Scalar, vector, and matrix mathematics: theory, facts, and formulas-revised and expanded edition*. Princeton Univ. Press, 2018.
- [33] J. Liu and P. Seiler, "Robust regret optimal control," *Int. Journal of Robust and Nonlinear Control*, vol. 34, no. 7, pp. 4532–4553, 2024.
- [34] S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge Univ. Press, 2018.