Measurement-Feedback Control with Optimal Data-Dependent Regret

Gautam Goel and Babak Hassibi

Abstract—Inspired by online learning, data-dependent regret has recently been proposed as a criterion for controller design. In the regret-optimal control paradigm, causal controllers are designed to minimize regret against a hypothetical optimal noncausal controller, which selects the globally cost-minimizing sequence of control actions given noncausal access to the disturbance sequence. We extend regret-optimal control to the more challenging measurement-feedback setting, where the online controller must compete against the optimal noncausal controller without directly observing the state or the driving disturbance. We show that no measurement-feedback controller can have bounded competitive ratio or regret which is bounded by the pathlength of the measurement disturbance. We do derive, however, a controller whose regret has optimal dependence on the joint energy of the driving and measurement disturbances, and another controller whose regret has optimal dependence on the pathlength of the driving disturbance and the energy of the measurement disturbance. The key technique we introduce is a reduction from regret-optimal measurement-feedback control to $H_\infty$-optimal measurement-feedback control in a synthetic system. We present numerical simulations which illustrate the efficacy of our proposed control algorithms.

I. INTRODUCTION

Inspired by online learning, data-dependent regret has recently been proposed as a criterion for controller design [4] In the regret-optimal control paradigm, causal controllers are designed to minimize regret against a hypothetical optimal noncausal controller, which selects the globally cost-minimizing sequence of control actions given noncausal access to the disturbance sequence. Controllers with low regret retain a performance guarantee relative to this strong benchmark irrespective of how the disturbance is generated; it is this universality which makes regret-optimal controllers an attractive alternative to traditional $H_2$ and $H_\infty$ controllers, which instead posit a specific disturbance-generating mechanism. The regret of the causal controller is bounded by some measure of the complexity of the disturbance sequence; several different complexity measures have been proposed, including the energy of the disturbance sequence, which measures the size of the disturbance, and the pathlength of the disturbance, which measures its variation over time. The alternative metric of competitive ratio, which is the worst-case ratio between the cost incurred by the causal controller and the cost incurred by the optimal noncausal controller, has also been proposed as a performance objective [11]. This metric can also be viewed as a special case of data-dependent regret, where the complexity measure is simply the offline optimal cost.

In a series of papers, Goel and Hassibi showed that for each of these complexity measures, the corresponding full-information controller can be computed via a reduction to full-information $H_\infty$ control [5, 6, 7]. In this paper we consider the more challenging problem of regret-optimal measurement-feedback control. In this setting, the causal controller is unable to directly observe the state or disturbance, and instead only has access to noisy linear observations of the state. As a result, the controller must not only choose which control actions to select, but must maintain and update an estimate of the state as new measurements arrive. In the $H_2$ setting it is well-known that the measurement-feedback problem can be decomposed into a stochastic full-information optimal control problem and a stochastic filtering problem; the fact that stochastic control algorithm and the filtering algorithm can be designed separately is usually called the Separation Principle. The $H_\infty$-optimal measurement-feedback controller can also be reduced to a filtering and full-information control problem, but the filtering and control algorithms are strongly coupled [12]. Intuitively, this is because an adversary which can optimize over the driving disturbance and the measurement noise simultaneously is strictly more powerful than one which must choose the disturbances separately.

We present the first characterization of measurement-feedback controllers with optimal regret against the optimal noncausal controller. We show that general linear systems do not admit measurement-feedback controllers with bounded competitive ratio or regret which is bounded by the pathlength of the measurement disturbance. We do, however, derive two new regret-optimal measurement-feedback controllers, the first of which attains regret with optimal dependence on the joint energy of the driving disturbance and the measurement disturbance, and the second of which attains regret with optimal dependence on the pathlength of the driving disturbance.
disturbance and the energy of the measurement disturbance. In both cases we reduce the problem of finding a controller with optimal regret in the original system to the problem of finding an $H_\infty$-optimal measurement-feedback controller in a synthetic system. We present numerical simulations which demonstrate that our regret-optimal controllers are competitive with standard $H_2$ and $H_\infty$ controllers.

II. RELATED WORK

The problem of obtaining a controller whose regret has optimal dependence on the energy of the disturbance was studied in [7] [15]. Recently, [3] and [14] described connections between regret-optimal control and the System-Level-Synthesis (SLS) framework [2], and used these connections to obtain regret-minimizing controllers that satisfy safety constraints. We also note the paper [6], which obtained a controller whose regret has optimal dependence on the pathlength of the disturbance. The problem of designing a controller with bounded competitive ratio was first studied in [10], and those results were extended in [11] [16]. A controller with optimal competitive ratio was obtained in [3]. We also note a parallel line of work [1] [17] which studies control through the weaker metric of policy regret; this metric compares the performance of the online controller to the best controller selected in hindsight from a parametric class of controllers, such as the class of state-feedback controllers. We note that this benchmark is strictly weaker than the the optimal noncausal controller we consider in this paper, as established in [9].

All of these aforementioned works focus on the full-information control setting; [8] was the first to study dynamic regret minimization in the measurement-feedback setting. In this paper, we focus on obtaining a measurement-feedback controller which minimizes regret against a clairvoyant, globally-optimal noncausal controller which observes the actual driving disturbance without any measurement noise; the paper [8] instead considers the problem of minimizing regret against a weaker $H_2$-optimal noncausal benchmark.

III. PRELIMINARIES

We consider a discrete-time LTI system that evolves according to the dynamics

$$x_{t+1} = Ax_t + Bu_t + Bu_t w_t,$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^m$ is the control input, $w_t \in \mathbb{R}^p$ is an exogenous disturbance, and $A, B_u, B_w$ are matrices of compatible dimensions. In each timestep, the controller receives a noisy linear measurement $y_t$ of the current state, where $y_t = Cx_t + v_t$ and $C \in \mathbb{R}^{r \times n}$ and $v_t \in \mathbb{R}^r$ is a measurement disturbance. The controller also incurs a quadratic cost

$$x_t^* Q x_t + u_t^* R u_t$$

in each timestep; it is convenient to assume that the dynamics are scaled so that $R = I$. Our goal is to design a controller which, given the measurements $y$, selects a control action $u$ so as to minimize the dynamic regret against a clairvoyant offline optimal controller which can observe the full sequence of disturbances $w$ ahead of time. Define $L = Q^{1/2}$, $s_t = Lx_t$. The dynamics and observation model can be neatly expressed

$$s = Fu + Gw, \ y = Hu + Jw + v,$$

where $F$ and $G$ are the strictly causal transfer operators mapping $u$ and $w$ to $s$ and $H$ and $J$ are the strictly causal transfer operators mapping $u$ and $w$ to $w$. We restrict our attention to control policies which are a causal linear function of the observations $y$, i.e., policies which set $u = Ky$ for some causal operator $K$. Solving for $y$, we see that

$$y = (I - HK)^{-1} (Jw + v),$$

implying that

$$u = K(I - HK)^{-1} (Jw + v).$$

We introduce the Youla parameterization $Q = K(I - HK)^{-1}$; we can easily recover $K$ from $Q$ by setting $K = (I + QH)^{-1} Q$. To each $K$, we associate the transfer operator

$$T_K : \begin{bmatrix} w \\ u \end{bmatrix} \rightarrow \begin{bmatrix} s \\ u \end{bmatrix}$$

given by

$$T_K = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} F \\ I \end{bmatrix} Q \begin{bmatrix} J & I \end{bmatrix}.$$

(1)

Recall that the optimal noncausal controller selects the control $u = K_0 w$, where

$$K_0 = -(I + F^* F)^{-1} F^* G.$$

The transfer operator associated to the optimal noncausal controller is therefore

$$T_{K_0} = \begin{bmatrix} -F(I + F^* F)^{-1} F^* G + G & 0 \\ -(I + F^* F)^{-1} F^* G & 0 \end{bmatrix}.$$ 

The zeros in the second column represent the fact that the optimal noncausal controller observes the actual disturbance $w$ and the control signal it selects is not at all affected by the measurement noise $v$; it is this disparity between the information available to the online and offline controllers which makes regret-optimal measurement-feedback control considerably more challenging than regret-optimal full-information control.
The regret of a causal controller $K$ against the optimal noncausal controller $K_0$ on the instance $(w, v)$ is simply the difference in their costs:

$$\left[ \begin{array}{c} w \\ v \end{array} \right]^* (T_K^* T_K - T_{K_0}^* T_{K_0}) \left[ \begin{array}{c} w \\ v \end{array} \right].$$

Similarly, the competitive ratio of $K$ is the worst-case ratio of their costs:

$$\sup_{w, v} \left[ \begin{array}{c} w \\ v \end{array} \right]^* T_K^* T_K \left[ \begin{array}{c} w \\ v \end{array} \right] / \left[ \begin{array}{c} w \\ v \end{array} \right]^* T_{K_0}^* T_{K_0} \left[ \begin{array}{c} w \\ v \end{array} \right].$$

Recall that the energy of a signal $w$ is $\|w\|^2$, while the pathlength of $w$ is $\sum_{i=-\infty}^{\infty} \|w_i - w_{i-1}\|^2$. The goal of this paper is to completely characterize the conditions under which it is possible to obtain a controller with bounded competitive ratio, or regret bounded by the energy or pathlength of $w$ and $v$. In those cases where it is possible to obtain a controller with bounded regret, we derive controllers with the tightest possible regret bounds.

IV. NON-EXISTENCE RESULTS

We first establish that in unstable systems there is no controller with bounded competitive ratio; in stable systems the only such controller is the controller which always sets $u = 0$. The key observation is that the competitive ratio bound implies that the online controller must incur zero cost on every disturbance sequences on which the optimal noncausal controller incurs zero cost. In particular, in order to have a bounded competitive ratio bound implies that the online controller must incur zero cost on every disturbance sequences on which the optimal noncausal controller incurs zero cost. This implies that the online controller must always set $u = 0$. In the measurement-feedback setting the controller is unable to directly observe $w$, and hence must set $u = 0$ at all times. The only setting under which this controller can be competitive is when the system is stable.

**Theorem 1.** Fix $\gamma > 0$ and suppose $A$ is unstable. There does not exist a causal measurement-feedback controller $\pi$ such that

$$J(\pi, w, v) < \gamma^2 J(\pi_0, w)$$

for all driving disturbances $w$ and all measurement disturbances $v$. If $A$ is stable, then the only competitive controller is the controller which always sets $u = 0$. This controller has competitive ratio $1 + \|F\|^2$.

**Proof.** Suppose there exists a measurement $y_0$ such that $\pi(y_0) \neq 0$. Set $w = 0$, $v = y_0$. The online controller incurs positive cost, but the offline optimal cost is zero, contradicting (2). Now suppose that the controller always sets $u = 0$, irrespective of $y_0$. In this case, the competitive ratio is simply

$$\sup_{w \in \ell_2} ||Gw||^2 / ||\Delta^{-1}Gw||^2,$$

where $\Delta$ is defined in Lemma 1. Let $w_2 = \Delta^{-1}Gw$. The competitive ratio becomes

$$\sup_{w_2 \in \ell_2} \frac{||\Delta w_2||^2}{||w_2||^2} = \sup_{w_2 \in \ell_2} \frac{||\Delta^* w_2||^2}{||w_2||^2} = \frac{w_2^*(I + FF^*)w_2}{||w_2||^2} = 1 + \sup_{w_2 \in \ell_2} \frac{w_2^*Fw_2}{||w_2||^2} = 1 + ||F||^2.$$

It is easy to check that $F$ is bounded if and only if $A$ is stable.

We next prove that there is no controller whose regret bounded by the joint pathlength of the driving disturbance and the measurement disturbance; this non-existence result holds for all linear systems, stable or unstable. The key observation is that the regret bound implies that the online controller must incur zero cost whenever $w$ and $v$ are constant. In the measurement-feedback setting the controller is unable to directly observe $w$ and $v$, and hence must set $u = 0$ at all times. However, there exist choices of $w$ and $v$ with zero pathlength such that this controller incurs positive regret, contradicting the pathlength bound.

**Theorem 2.** Fix $\gamma > 0$. There does not exist a causal controller $\pi$ such that

$$J(\pi, w, v) - J(\pi_0, w) < \gamma^2 (||D_p w||^2 + ||D_r v||^2)$$

for all driving disturbances $w$ and all measurement disturbances $v$.

**Proof.** Suppose there exists a measurement $y_0$ such that $\pi(y_0) \neq 0$. Set $w = 0$, $v = y_0$. The online controller incurs positive cost, but the offline optimal cost is zero, hence the online controller incurs positive regret. This contradicts the fact that the right-hand side of (3) is zero. Now suppose that the controller always sets $u = 0$, irrespective of $y_0$. The left-hand side of (3) is

$$||Gw||^2 - ||\Delta^{-1}Gw||^2 = w^*G^* (I - (I + FF^*)^{-1}) Gw = w^*G^* F(I + FF^*)^{-1} F^* Gw,$$

where $\Delta$ is defined in Lemma 1. Set $w$ to be a nonzero constant sequence such that $F^* Gw \neq 0$. Set $v = 0$. Then the left-hand side of (3) is strictly positive, while the right-hand side is zero for all $\gamma$. We now prove that in unstable systems there is no controller whose regret is bounded by the energy of
the driving disturbance and the pathlength of the mea-
urement disturbance; in stable systems the only such
controller is the controller which always sets \( u = 0 \). The
key observation is that the regret bound implies that the
online controller must incur zero cost whenever \( w = 0 \)
and \( v \) is constant. In the measurement-feedback setting
the controller is unable to directly observe \( w \) and \( v \), and
hence must set \( u = 0 \) at all times.

**Theorem 3.** Fix \( \gamma > 0 \) and suppose \( A \) is unstable. There
does not exist a causal controller \( \pi \) such that

\[
J(\pi, w, v) - J(\pi_0, w) < \gamma^2 \left( \|w\|_2^2 + \|D_r v\|_2^2 \right)
\]

(4)

for all driving disturbances \( w \) and all measurement
disturbances \( v \). If \( A \) is stable, then the only controller
which satisfies (4) for any value of \( \gamma \) is the "zero
controller," which always sets \( u = 0 \). This controller
satisfies (4) with \( \gamma = \|\Delta_2^{-1} F^* G\| \), where \( \Delta_2 \) is the
unique causal and causally invertible operator such that
\( \Delta_2 \Delta_2 = I + F^* F \).

**Proof.** Suppose there exists a measurement \( y_0 \) such that
\( \pi(y_0) \neq 0 \). Set \( w = 0, v = y_0 \). The online controller
incurs positive cost, but the offline optimal cost is zero,
and the online controller incurs positive regret. This
contradicts the fact that the right-hand side of (4) is
zero. Now suppose that the controller always sets \( u = 0 \),
irrespective of \( y_0 \). The left-hand side of (4) is

\[
\|Gw\|^2 - \|\Delta_2^{-1} Gw\|^2 = w^* G^* (I - (I + F^* F)^{-1}) Gw
\]

\[
= w^* G^* F (I + F^* F)^{-1} F^* G w,
\]

where \( \Delta \) is defined in Lemma 1. The smallest possible
value of \( \gamma^2 \) such that (4) holds is given by

\[
\sup_{w \in L^2} \frac{w^* G^* F (I + F^* F)^{-1} F^* G w}{\|w\|^2} = \|\Delta_2^{-1} F^* G\|^2,
\]

where \( \Delta_2 \) is the unique causal and causally invertible
operator such that \( \Delta_2 \Delta_2 = I + F^* F \). We note that
the operator \( \Delta_2^{-1} F^* G \) is bounded if and only if \( A \) is
stable. \( \square \)

V. REGRET BOUNDED BY THE JOINT ENERGY OF \( w \)
AND \( v \)

In this section we describe a causal controller \( \pi \)
whose regret is bounded by the joint energy of the
driving disturbance and the measurement disturbance.
This presents a sharp contrast with the negative results
of the previous sections; the key difference is that the
only pair of disturbances \((w, v)\) whose joint energy
is zero is simply \( w = 0, v = 0 \). It is easy to guarantee
that \( K \) sets \( u = 0 \) on this specific instance without
also requiring that \( K \) sets \( u = 0 \) on all other instances.
We derive the measurement-feedback controller whose
regret has optimal dependence on the joint energy of the
driving disturbance and the measurement disturbance via
a reduction to \( H_\infty \) measurement-feedback control.

**Theorem 4.** Fix \( \gamma > 0 \) and define \( \hat{A}, \hat{B}_u, \hat{B}_w, \hat{C}, \hat{L} \)
as in (10). There exists a causal measurement-feedback
controller \( \pi \) such that

\[
J(\pi, w, v) - J(\pi_0, w) < \gamma^2 \left( \|w\|_2^2 + \|v\|_2^2 \right)
\]

(5)

if and only if the control DARE

\[
P_c = \hat{A}^* P_c \hat{A} + \hat{L}^* \hat{L} - K_c^* R_c K_c,
\]

and the estimation DARE

\[
P_c = \hat{A} P_c \hat{A}^* + \hat{B}_w \hat{B}_w^* - K_c R_c K_c^*,
\]

where we define

\[
\begin{align*}
K_c &= R_c^{-1} \begin{bmatrix} \hat{B}_u^* \\ \hat{B}_w^* \end{bmatrix} P_c \hat{A} \\
R_c &= \begin{bmatrix} I_m & 0 \\ 0 & -I_p \end{bmatrix} + \begin{bmatrix} \hat{B}_u^* \\ \hat{B}_w^* \end{bmatrix} P_c \begin{bmatrix} \hat{B}_u & \hat{B}_w \end{bmatrix}, \\
K_c &= \hat{A} P d \begin{bmatrix} \hat{C} \\ \hat{L} \end{bmatrix} R_c^{-1}, \\
R_c &= \begin{bmatrix} I_r & 0 \\ 0 & -I_n \end{bmatrix} + \begin{bmatrix} \hat{C} \\ \hat{L} \end{bmatrix} P_c \begin{bmatrix} \hat{C}^* & \hat{L}^* \end{bmatrix},
\end{align*}
\]

have solutions \( P_c \geq 0 \) and \( P_c \geq 0 \) such that

1. The matrix \( \hat{A} - \begin{bmatrix} \hat{B}_u & \hat{B}_w \end{bmatrix} K_c \) is stable.
2. The matrix \( R_c \) has \( m \) positive eigenvalues and \( n \)
   negative eigenvalues.
3. The matrix \( \hat{A} - K_c \begin{bmatrix} \hat{C} \\ \hat{L} \end{bmatrix} \) is stable.
4. The matrix \( R_c \) has \( m \) positive eigenvalues and \( n \)
   negative eigenvalues.
5. \( \rho(P_c P_c^*) < 1 \).

If these conditions are satisfied, then one possible choice
of \( \pi \) is given by

\[
u_t = -K_u (\xi_t + P \hat{C}^* (I_r + \hat{C} P \hat{C}^*)^{-1} (y_t - \hat{C} \xi_t)),
\]

where the synthetic state \( \xi \in \mathbb{R}^{2n} \) evolves according to
the linear dynamics equation

\[
\xi_{t+1} = (\hat{A} - \hat{B}_w K_w) (\xi_t + P \hat{C}^* (I_r + \hat{C} P \hat{C}^*)^{-1} (y_t - \hat{C} \xi_t)) + \hat{B}_u u_t.
\]

The matrices \( K_u \in \mathbb{R}^{m \times 2n} \) and \( K_w \in \mathbb{R}^{p \times 2n} \) are
defined as

\[
\begin{bmatrix} K_u \\ K_w \end{bmatrix} = K_c,
\]

and we define \( P \in \mathbb{R}^{2n \times 2n} \) as

\[
P = P_c (I_r - P_c \hat{C}_c)^{-1}.
\]
We note that the regret-optimal controller can be easily obtained via bisection on $\gamma$.

Proof. The regret condition (5) can be rewritten as
\[
T_K^*T_K \preceq \left( \gamma^2 \begin{bmatrix} I_p & 0 \\ 0 & I_r \end{bmatrix} + T_{K_0}^*T_{K_0} \right).
\] (6)

Let $\Delta_2$ be the unique causal and causally invertible operator such that
\[
\gamma^2 I_p + G^* (I + FF^*)^{-1} G = \Delta_2^* \Delta_2.
\] (7)

Then
\[
\gamma^2 \begin{bmatrix} I_p & 0 \\ 0 & I_r \end{bmatrix} + T_{K_0}^*T_{K_0} = \begin{bmatrix} \Delta_2 & 0 \\ 0 & \gamma I \end{bmatrix}^* \begin{bmatrix} \Delta_2 & 0 \\ 0 & \gamma I \end{bmatrix}.
\]

Define
\[
\begin{bmatrix} \hat{w} \\ \hat{v} \end{bmatrix} = \begin{bmatrix} \Delta_2 & 0 \\ 0 & \gamma I \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix}.
\]

Condition (6) can be rewritten as
\[
\begin{bmatrix} \hat{w} \\ \hat{v} \end{bmatrix}^* \begin{bmatrix} \hat{w} \\ \hat{v} \end{bmatrix} < \left\| \begin{bmatrix} \hat{w} \\ \hat{v} \end{bmatrix} \right\|^2,
\]

or equivalently as
\[
\left\| T_K \right\| < 1,
\]
where we define
\[
T_K = \begin{bmatrix} \Delta_2 & 0 \\ 0 & \gamma I \end{bmatrix}^{-1}.
\]

Using the parameterization (1) of $T_K$, we see that
\[
T_K = \left[ \begin{array}{cc} G & 0 \\ -1 & -1 \end{array} \right] + \begin{bmatrix} F & 0 \\ I & -1 \end{bmatrix} (F^{-1} Q) \left[ \begin{array}{cc} G & 0 \\ -1 & -1 \end{array} \right] \left[ \begin{array}{cc} G & 0 \\ -1 & -1 \end{array} \right] -1.
\]

Notice that $T_K$ itself has the form of a transfer operator described in (1); it is the transfer operator with Youla parameter $\hat{Q} = \gamma^{-1} Q$ in the system
\[
s = \hat{F} u + \hat{G} \hat{w}, \quad \hat{y} = \hat{H} u + \hat{w},
\]
where we define $\hat{F} = F, \hat{G} = G \Delta_2^{-1}, \hat{H} = \gamma H, \hat{w} = \gamma J \Delta_2^{-1}$. It is now clear that a controller $K$ satisfying (3) exists if and only if there exists a controller $\hat{K}$ in the system (8) such that $\left\| T_{\hat{K}} \right\| < 1$. If such a controller $\hat{K}$ exists, then we can easily recover $K$ from $\hat{K}$ by setting $K = \gamma \hat{K}$; notice that this is the unique choice of $K$ which is consistent with the relations $\hat{H} = \gamma H, \hat{Q} = \gamma^{-1} Q$.

In order to assign state-space structure to $\hat{F}, \hat{G}, \hat{H}, \hat{J}$, we must first find $\Delta_2(z)$. Let $\Delta$ be the unique causal and causally invertible operator such that
\[
\Delta F = \Delta \Delta^*.
\]

Lemma [1] shows that
\[
\Delta(z) = (I + L(zI - A)^{-1} K) \Sigma^{1/2}.
\]

It follows that $\Delta^{-1}(z) G(z)$ is given by
\[
\Delta^{-1}(z) G(z) = \Sigma^{-1/2} L(zI - (A - KL))^{-1} B_w.
\]

Define $\tilde{A} = A - KL$. Notice that $\gamma^2 I_p + G^* \Delta^* \Delta^{-1} G$ can be rewritten as
\[
\begin{bmatrix} B_w^*(z^{-1} I_n - \tilde{A})^{-1} & \tilde{A} \end{bmatrix} \begin{bmatrix} L^* \Sigma^{-1} L & 0 \\ 0 & \gamma^2 I_p \end{bmatrix} \begin{bmatrix} I_n - \tilde{A} \end{bmatrix}^{-1} B_w.
\]

Applying Lemma [2] we see that this equals
\[
\begin{bmatrix} B_w^*(z^{-1} I_n - \tilde{A})^{-1} & \tilde{A} \end{bmatrix} \begin{bmatrix} (I_n - \tilde{A})^{-1} B_w \end{bmatrix},
\]

where $P_2$ is an arbitrary Hermitian matrix and we define
\[
\Lambda_2(P_2) = \begin{bmatrix} L^* \Sigma^{-1} L + P_2 - \tilde{A}^* P_2 \tilde{A} & \tilde{A}^* P_2 B_w \\ B_w^* P_2 \tilde{A} & \gamma^2 I + B_w^* P_2 B_w \end{bmatrix}.
\]

Notice that the $\Lambda_2(P_2)$ can be factored as
\[
\begin{bmatrix} I & K_2^*(P_2) \\ 0 & I \end{bmatrix} \begin{bmatrix} \Gamma_2(P_2) & 0 \\ 0 & \Sigma_2(P_2) \end{bmatrix} \begin{bmatrix} I & 0 \\ K_2(P_2) & I \end{bmatrix},
\]

where we define
\[
\Gamma_2(P_2) = L^* \Sigma^{-1} L + P_2 + \tilde{A}^* P_2 \tilde{A} - K_2^*(P_2) \Sigma_2 K_2(P_2), \Sigma_2(P_2) = \sigma_2^{-1}(P_2) B_w^* P_2 \tilde{A}, \Sigma_2(P_2) = \gamma^2 I + B_w^* P_2 B_w.
\]

It is clear that $(\tilde{A}, B_w)$ is stabilizable (in fact, $\tilde{A}$ is stable), therefore the Riccati equation $\Gamma_2(p_2) = 0$ has a unique stabilizing solution (Theorem E.6.2 in [13]). Suppose $P_2$ is chosen to be this solution, and define $K_2 = K_2(P_2), \Sigma_2 = \Sigma_2(P_2)$. We immediately obtain the factorization (7), where we define
\[
\Delta_2(z) = \Sigma_2^{-1/2} (I + K_2(zI - \tilde{A})^{-1} B_w).
\] (9)

We have
\[
\Delta_2^{-1}(z) = (I - K_2(zI - (A - B_w K_2))^{-1} B_w) \Sigma_2^{-1/2}.
\]

We note that $\tilde{A} - B_w K_2$ is stable and hence $\Delta^{-1}(z)$ is causal and bounded since its poles are strictly contained in the unit circle. Define
\[
\begin{align*}
\hat{A} &= \begin{bmatrix} A & -B_w K_2 \\ 0 & A - B_w K_2 \end{bmatrix}, \\
\hat{B}_w &= \begin{bmatrix} B_w \Sigma_2^{-1/2} \\ B_w \Sigma_2^{-1/2} \end{bmatrix}, \\
\hat{B}_u &= \begin{bmatrix} B_u \\ 0 \end{bmatrix}, \\
\hat{C} &= \begin{bmatrix} \gamma C & 0 \end{bmatrix}, \\
\hat{L} &= \begin{bmatrix} L & 0 \end{bmatrix}.
\end{align*}
\] (10)
Recall that $\tilde{F} = F, \tilde{G} = G \Delta_{\gamma}^{-1}, \tilde{H} = \gamma H, \tilde{J} = \gamma J \Delta_{\gamma}^{-1}$; it follows that $\tilde{F}, \tilde{G}, \tilde{H}, \tilde{J}$ are given by

$$\tilde{F}(z) = \tilde{L}(z - \hat{A})^{-1} \hat{B}_u,$$

$$\tilde{G}(z) = \tilde{L}(zI - \hat{A}) \hat{B}_w,$$

$$\tilde{H}(z) = \tilde{C}(zI - \hat{A})^{-1} \hat{B}_u,$$

$$\tilde{J}(z) = \tilde{C}(zI - \hat{A}) \hat{B}_w.$$
Using the parameterization (1) of $T_K$, we see that

$$T^*_K = \begin{bmatrix} G\Delta^{-1} & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} F \end{bmatrix} \left( \gamma^{-1}Q \right) \left[ \gamma J \Delta^{-1} \right].$$

Notice that $T^*_K$ itself has the form of a transfer operator described in (1); it is the transfer operator with Youla parameter $Q = \gamma^{-1}Q$ in the system

$$s = \tilde{F} u + \tilde{G} \tilde{w}, \quad \tilde{y} = \tilde{H} u + \tilde{J} \tilde{w} + \tilde{v}, \quad (14)$$

where we define $\tilde{F} = F, \tilde{G} = G\Delta^{-1}, \tilde{H} = \gamma H, \tilde{J} = \gamma J \Delta^{-1}$. It is now clear that a controller $K$ satisfying (12) exists if and only if there exists a controller $\tilde{K}$ in the system (14) such that $\|T^*_K\| < 1$. If such a controller $\tilde{K}$ exists, then we can easily recover $K$ from $\tilde{K}$ by setting $K = \gamma \tilde{K}$; notice that this is the unique choice of $K$ which is consistent with the relations $\tilde{H} = \gamma H, Q = \gamma^{-1}Q$.

In order to assign state-space structure to $\tilde{F}, \tilde{G}, \tilde{H}, \tilde{J}$, we must first find $\Delta_2(z)$. Let $\Delta$ be the unique causal and causally invertible operator such that

$$I + FF^* = \Delta\Delta^*.$$

Lemma 1 shows that

$$\Delta(z) = (I + L(zI - A)^{-1}K)\Sigma^{1/2}.$$

It follows that $\Delta^{-1}(z)G(z)$ is given by

$$\Delta^{-1}(z)G(z) = \Sigma^{-1/2}L(zI - (A - KL))^{-1}B_w.$$

Define

$$\begin{align*}
\tilde{L} &= \gamma^{-1/2}L \\
\tilde{A} &= A - KL \\
\tilde{B}_w &= B_w
\end{align*}$$

Notice that we can rewrite

$$\begin{bmatrix} \gamma^2 D_p^*(-^*)D_p(z) + G^*\Delta^*\Delta^{-1}G \\ 0 \\ 0 \\ \gamma^2 I \end{bmatrix}$$

as

$$\begin{bmatrix} \tilde{B}_w(z^*-^*) \tilde{L} \tilde{S} \gamma I \\ \tilde{S}^* \gamma^2 I \\ (zI - \tilde{A})^{-1} \tilde{B}_w \end{bmatrix}\begin{bmatrix} \tilde{L} \tilde{S} \gamma I \\ \tilde{S}^* \gamma^2 I \\ (zI - \tilde{A})^{-1} \tilde{B}_w \end{bmatrix}.$$

Applying Lemma 2, we see that this equals

$$\begin{bmatrix} \tilde{B}_w(z^*-^*) \tilde{L} \tilde{S} \gamma I \\ 0 \end{bmatrix} \Lambda_2(P_2) \begin{bmatrix} (zI - \tilde{A})^{-1} \tilde{B}_w \\ I \end{bmatrix},$$

where $P_2$ is an arbitrary Hermitian matrix and we define

$$\Lambda_2(P_2) = \begin{bmatrix} \tilde{L} + L^* - P_2 + \tilde{A}P_2 \tilde{S}^* + \tilde{A}P_2 \tilde{B}_w \\ 0 \\ 0 \\ \gamma^2 I + \tilde{B}_w \end{bmatrix}.$$
close to $(0, 0)$ as possible. The discrete-time dynamics are

\[
\begin{bmatrix}
  x_{t+1} \\
  \dot{x}_{t+1}
\end{bmatrix} =
\begin{bmatrix}
  1 & \delta_t \\
  0 & 1
\end{bmatrix}
\begin{bmatrix}
  x_t \\
  \dot{x}_t
\end{bmatrix} +
\begin{bmatrix}
  0 \\
  \delta_t
\end{bmatrix} u_t +
\begin{bmatrix}
  0 \\
  \delta_t
\end{bmatrix} w_t,
\]

where $\delta_t$ is the discretization parameter; in our experiments we set $\delta_t = 0.1$ seconds. In our experiments we take $Q = I_2$ and initialize $x$ and $\dot{x}$ to zero. We study the relative performance of the $H_2$-optimal, $H_\infty$-optimal, energy-optimal, and pathlength-optimal controllers across various kinds of driving disturbances; in all of our experiments we take the measurement disturbance $v$ to be sampled i.i.d. from $\mathcal{N}(0, 1)$ in each timestep.

We first consider an optimistic setting where both the driving disturbance $w$ and the measurement disturbance $v$ are drawn i.i.d from $\mathcal{N}(0, 1)$ in each timestep (Figure 1). As expected, the $H_2$ controller performs the best out of all causal controllers, with the energy-optimal controller in second place. The poor performance of the pathlength-optimal controller is explained by the fact that a Gaussian driving disturbance has high pathlength relative to its energy.

We next consider an adversarial setting where the driving disturbance is an impulse, i.e. the disturbance is zero across the whole time horizon except for a spike near $t = 500$ (Figure 2). As expected, the $H_\infty$ controller performs the best out of all causal controllers, with the energy-optimal controller close behind. The overly optimistic $H_2$ controller performs poorly on this disturbance.

Lastly, we consider a setting where the driving disturbance is a Gaussian random walk, i.e. $w_t$ is the sum of the first $t$ of a set of random variables, each of which is drawn i.i.d from $\mathcal{N}(0, 1)$. Notice that the energy of this disturbance is much higher than its pathlength. As expected, the pathlength-optimal controller now outperforms all other causal controllers and closely tracks the clairvoyant noncausal controller.

REFERENCES

[1] Naman Agarwal et al. “Online control with adversarial disturbances”. In: arXiv preprint arXiv:1902.08721 (2019).
[2] James Anderson et al. “System level synthesis”. In: Annual Reviews in Control 47 (2019), pp. 364–393.
[3] Alexandre Didier, Jerome Sieber, and Melanie N Zeilinger. “A system level approach to regret optimal control”. In: IEEE Control Systems Letters (2022).
[4] Gautam Goel. “Regret-Optimal Control”. PhD thesis. California Institute of Technology, 2022.
[5] Gautam Goel and Babak Hassibi. “Competitive Control”. In: arXiv preprint arXiv:2107.13657 (2021).
\[ K = APL\Sigma^{-1}, \quad \Sigma = I + LPL, \]

and \( P \) is the unique Hermitian solution to the Riccati equation
\[ P = B_uB_u^* + APA^* - APL(I + LPL)^{-1}LPA^*. \]

**Proof.** We expand \( I + F(z)F(z^-)^* \) as
\[
\begin{bmatrix}
L(zI - A)^{-1} & I \\
0 & I
\end{bmatrix}
\begin{bmatrix}
B_uB_u^* & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
(z^-I - A)^{-*}L \\
I
\end{bmatrix}.
\]

Applying Lemma 3, we see that this equals
\[
[L(zI - A)^{-1} I] \Lambda(P) \begin{bmatrix}
(z^-I - A)^{-*}L \\
I
\end{bmatrix},
\]

where \( P \) is an arbitrary Hermitian matrix and we define
\[
\Lambda(P) = \begin{bmatrix}
B_uB_u^* - P + APA^* & APL \\
LPA^* & L + LPL
\end{bmatrix}.
\]

Notice that the \( \Lambda(P) \) can be factored as
\[
\begin{bmatrix}
I & K(P) \\
0 & I
\end{bmatrix} \begin{bmatrix}
\Gamma(P) & 0 \\
0 & \Sigma(P)
\end{bmatrix} \begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix},
\]

where we define
\[
\Gamma(P) = B_uB_u^* - P + APA^* - K(P)\Sigma(P)K^*(P),
\]
\[
K(P) = APL\Sigma(P)^{-1},
\]
\[
\Sigma(P) = I + LPL.
\]

By assumption, \((A, B_u)\) is stabilizable and \((A, L)\) is detectable, therefore the Riccati equation \( \Gamma(P) = 0 \) has a unique stabilizing solution (see, e.g. Theorem E.6.2 in [13]). Suppose \( P \) is chosen to be this solution, and define \( K = K(P), \Sigma = \Sigma(P) \). We immediately obtain the canonical factorization
\[
I + F(z)F(z^-)^* = \Delta(z)\Delta^*(z^-),
\]
where we define
\[
\Delta(z) = (I + L(zI - A)^{-1}K)\Sigma^{1/2}. \tag{19}
\]

**Lemma 2.** For all \( H, F \) and all Hermitian matrices \( P \), we have
\[
[H^*(z^{-1}I - F^*)^{-1} I] \Omega(P) \begin{bmatrix}
(zI - F)^{-1}H \\
I
\end{bmatrix} = 0,
\]
where we define
\[
\Omega(P) = \begin{bmatrix}
-P + F^*PF & F^*PH \\
H^*PF & H^*PH
\end{bmatrix}.
\]

**Proof.** This identity is essentially the “transpose” of Lemma 3 and is easily verified via direct calculation. \( \square \)

### VIII. APPENDIX

#### A. Missing Lemmas

We present proofs of some of the key lemmas used in the main body.

**Lemma 1.** The following canonical factorization holds:
\[
I + F(z)F(z^-)^* = \Delta(z)\Delta^*(z^-),
\]
where we define
\[
\Delta(z) = (I + L(zI - A)^{-1}K)\Sigma^{1/2}. \tag{18}
\]
Lemma 3. For all $H, F$ and all Hermitian matrices $P$, we have
\[
[H(zI - F)^{-1} I] \Omega(P) \left[ (z^{-1}I - F^*)^{-1}H^* \right] = 0,
\]
where we define
\[
\Omega(P) = \begin{bmatrix}
-P + FPF^* & FP^*

HPF^* & HPH^*
\end{bmatrix}.
\]

Proof. This identity is a special case of Lemma 4 in it also appears as Lemma 12.3.3 in “Indefinite-Quadratic Estimation and Control” by Hassibi, Sayed, and Kailath.

Lemma 4. For all $H_1, H_2, F_1, F_2$ and all matrices $W$, we have
\[
[H_1(zI - F_1)^{-1} I] \Omega(W) \left[ (z^{-1}I - F_2^*)^{-1}H_2^* \right] = 0,
\]
where we define
\[
\Omega(W) = \begin{bmatrix}
-W + F_1WF_2^* & F_1WH_2^*

H_1WF_2^* & H_1WH_2^*
\end{bmatrix}.
\]

Proof. Notice that $\Omega(W)$ can be rewritten as
\[
\Omega(W) = \begin{bmatrix}
F_1 & W & F_2 & H_2

H_1 & 0 & 0 & 0
\end{bmatrix}.
\]
The proof is immediate after observing that
\[
[H_1(zI - F_1)^{-1} I] \begin{bmatrix}
F_1 \\
H_1
\end{bmatrix} = H_1(zI - F_1)^{-1}z,
\]
\[
\begin{bmatrix}
F_2 & H_2

F_1 & H_1
\end{bmatrix} \begin{bmatrix}
(z^{-1}I - F_2^*)^{-1}H_2^* \\
I
\end{bmatrix} = z^{-1}(z^{-1}I - F_2^*)^{-1}H_2^*.
\]

B. A state-space model for the $H_\infty$-optimal measurement-feedback controller

Theorem 6 (Theorem 13.3.5 in [12]). A causal measurement-feedback controller $K$ such that
\[
\|T_K\| < 1
\]
exists if and only if the control DARE
\[
P_c = A^*P_cA + L^*L - K_c^*R_cK_c
\]
and the estimation DARE
\[
P_e = AP_eA^* + B_wB_w^* - K_cR_cK_c^*,
\]
where we define
\[
K_e = R_e^{-1} \begin{bmatrix}
B_u^* \\
B_w
\end{bmatrix} P_e A
\]
\[
R_e = \begin{bmatrix} I \quad 0 \\
0 \quad -I_p \end{bmatrix} + \begin{bmatrix} B_u^* \\
B_w
\end{bmatrix} P_e \begin{bmatrix} B_u & B_w \end{bmatrix},
\]

have solutions $P_c \succeq 0$ and $P_e \succeq 0$ such that
1) The matrix $A - [B_u \quad B_w]K_e$ is stable.
2) The matrix $R_e$ has $m$ positive eigenvalues and $p$ negative eigenvalues.
3) The matrix $A - K_c \begin{bmatrix} C \\ L \end{bmatrix}$ is stable.
4) The matrix $R_e$ has $\gamma$ positive eigenvalues and $n$ negative eigenvalues.
5) $\rho(P_eP_e) < 1$.

If these conditions are satisfied, then one possible choice of $K$ is given by
\[
u_t = -K_u(\tilde{x}_t + PC^*(I_r + CPC^*)^{-1}(y_t - C\tilde{x}_t)),
\]
where the state-estimate $\tilde{x}_t$ is given by the recursion
\[
\tilde{x}_{t+1} = (A-B_wK_w)(\tilde{x}_t + PC^*(I_r + CPC^*)^{-1}(y_t - C\tilde{x}_t)) + B_uu_t,
\]
and $K_u \in \mathbb{R}^{m \times n}$ and $K_w \in \mathbb{R}^{p \times n}$ are defined as
\[
\begin{bmatrix} K_u \\
K_w
\end{bmatrix} = K_e,
\]
and we define $P \in \mathbb{R}^{n \times n}$ as
\[
P = P_e(I_n - P_eP_e)^{-1}.
\]

While Theorem 6 tells us how to determine whether there exists a controller $K$ such that $\|T_K\| < 1$, it does not directly answer the more general question if there exists a controller $K$ such that $\|T_K\| < \gamma$, for any fixed $\gamma > 0$. This condition is equivalent to $\|T_{\gamma K}\| < 1$, where we define $T_{\gamma K} = \gamma^{-1}T_K$. Notice that
\[
T_{\gamma K} = \begin{bmatrix} \gamma^{-1}G & 0 \\
0 & I
\end{bmatrix} + \begin{bmatrix} F \\
0
\end{bmatrix} (\gamma^{-1}Q)^{1/2} \begin{bmatrix} J & I
\end{bmatrix}.
\]

Recall that $Q = K(I - HK)^{-1}$, implying that $\gamma^{-1}Q = (\gamma^{-1}K)(I - (\gamma H)(\gamma^{-1}K))^{-1}$. It follows that the controller $K$ satisfying $\|T_K\| < \gamma$ in the system $F, G, J, H$ is precisely $\gamma K$, where $K$ is the controller satisfying $\|T_{\gamma K}\| < 1$ in the system $\tilde{F}, \tilde{G}, \tilde{J}, \tilde{H}$ and $\tilde{F} = F, \tilde{G} = G,$ 

...
\( \gamma^{-1}G, \hat{H} = \gamma H, \hat{J} = J \). We can assign state-space structure to \( \hat{F}, \hat{G}, \hat{H}, \hat{J} \) as follows. Recall that

\[
F(z) = L(zI - A)^{-1}B_u, \quad G(z) = L(zI - A)^{-1}B_w, \\
H(z) = C(zI - A)^{-1}B_u, \quad J(z) = C(zI - A)^{-1}B_w. 
\]

Define
\[
\hat{C} = \gamma C, \quad \hat{B}_w = \gamma^{-1}B_w. 
\]

We have
\[
\hat{F}(z) = L(zI - A)^{-1}\hat{B}_u, \quad \hat{G}(z) = L(zI - A)\hat{B}_w, \\
\hat{H}(z) = \hat{C}(zI - A)^{-1}B_u, \quad \hat{J}(z) = \hat{C}(zI - A)\hat{B}_w. 
\]

We can hence use Theorem 6 to check the existence of a controller \( K \) such that \( \|T_K\| < \gamma \), for any \( \gamma > 0 \).