A Distributionally Robust Approach to Regret Optimal Control using the Wasserstein Distance

Feras Al Taha, Shuhao Yan, Eilyan Bitar

Abstract— This paper proposes a distributionally robust approach to regret optimal control of discrete-time linear dynamical systems with quadratic costs subject to a stochastic additive disturbance on the state process. The underlying probability distribution of the disturbance process is unknown, but assumed to lie in a given ball of distributions defined in terms of the type-2 Wasserstein distance. In this framework, strictly causal linear disturbance feedback controllers are designed to minimize the worst-case expected regret. The regret incurred by a controller is defined as the difference between the cost it incurs in response to a realization of the disturbance process and the cost incurred by the optimal noncausal controller which has perfect knowledge of the disturbance process realization at the outset. Building on a well-established duality theory for optimal transport problems, we derive a reformulation of the minimax regret optimal control problem as a tractable semidefinite program. Using the equivalent dual reformulation, we characterize a worst-case distribution achieving the worst-case expected regret in relation to the distribution at the center of the Wasserstein ball. We compare the minimax regret optimal control design method with the distributionally robust optimal control approach using an illustrative example and numerical experiments.

I. INTRODUCTION

There is growing interest in utilizing regret-based performance metrics, originally proposed in [1], to design controllers for uncertain dynamical systems. Loosely speaking, regret measures the loss in performance incurred by a causal controller relative to a clairvoyant controller that has complete knowledge of the underlying system dynamics and disturbance process at the outset. The regret minimization paradigm has been used to address a variety of estimation and decision problems, including robust estimation [2], [3], multi-armed bandits [4], online convex optimization [5], and adaptive control [6].

In the context of linear quadratic (LQ) control problems with unknown state and input matrices, [7] and [8] design learning-enabled controllers that achieve expected regret rates which scale optimally with the control horizon. More recently, [9] and [10] introduced the framework of regret optimal control for LQ problems, where the disturbance acting on the system is assumed to be adversarial in nature. Using operator-theoretic techniques from robust control, controllers in state-space representation are synthesized to

minimize the worst-case regret over all possible disturbances with bounded energy. Building on this framework, [11] and [12] synthesize regret optimal controllers subject to state and input constraints. Using the system level parameterization [13], they provide equivalent reformulations of the resulting minimax regret optimal control problems as semidefinite programs (SDPs). While the resulting controllers are robust with respect to the worst-case disturbance realization, they may be conservative if the underlying disturbance is actually stochastic in nature.

Taking a probabilistic view, we propose an alternative approach to regret optimal control, where controllers are designed to minimize the worst-case expected regret over all probability distributions in a given ambiguity set specified in terms of the type-2 Wasserstein distance. In this setting, the expected regret is defined as the difference between an expected cost incurred by a causal controller and the expected cost incurred by the optimal noncausal controller with perfect knowledge of the disturbance trajectory at the outset. The proposed approach can also be thought of as an alternative to distributionally robust optimal (DRO) control, see, i.e., [14]–[16] and references therein. A potential drawback of the DRO control approach, which aims to minimize the worst-case expected cost across a given ambiguity set of distributions, is that resulting controllers may incur a large expected regret for certain distributions in the ambiguity set. In contrast, by minimizing the worst-case expected regret, the approach proposed in this paper ensures that resulting controllers have uniformly small regret for all distributions in the given ambiguity set. In Section III, we provide a motivating example that clearly illustrates some of the trade-offs inherent to these two methods.

Summary of contributions: In this paper, we propose a new framework for regret optimal controller synthesis using distributionally robust optimization methods. The control problem is formulated as a minimax optimization problem, involving a worst-case expectation problem over a Wasserstein ball of distributions. Utilizing duality for distributionally robust optimization problems [17], we obtain an equivalent convex reformulation of the worst-case expectation problem, which facilitates the reformulation of the minimax regret optimal control problem as a SDP. Our result extends a number of related results on the exact reformulation of worst-case expectation problems involving quadratic objective functions [18], [19]. We also provide an explicit characterization of a worst-case distribution achieving the worst-case expectation in relation to the central distribution in the ambiguity set.

The remainder of this paper is organized as follows. In Section II, we formulate the minimax regret optimal control problem for linear disturbance feedback control policies. In Section III, we provide a stylized example that clearly illustrates some of the strengths and limitations of our proposed approach. In Section IV, we derive an equivalent reformulation of the minimax regret optimal control problem as a SDP. In Section V, we present numerical experiments which show that the proposed approach can significantly outperform the distributionally robust optimal control approach in certain settings. We conclude the paper in Section VI.

Notation: Let $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of real numbers and nonnegative real numbers, respectively. Let $\mathbb{S}^n$ denote the set of all symmetric matrices in $\mathbb{R}^{n \times n}$. Denote the cone of $n \times n$ real symmetric positive definite (semidefinite) matrices by $\mathbb{S}^n_{++}$ ($\mathbb{S}^n_+$). Denote the Euclidean norm of a vector $x \in \mathbb{R}^n$ by $\|x\|$. Given matrices $A, B \in \mathbb{S}^n$, the relation $A \succ B$ means $A - B \in \mathbb{S}^n_{++}$, and the relation $A \succeq B$ means $A - B \in \mathbb{S}^n_+$. Denote the maximum and minimum eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$ by $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$, respectively. Given a matrix $A \in \mathbb{S}^n$, the matrix $A^{\frac{1}{2}}$ denotes the (symmetric) positive semidefinite square root of A. Let $\mathcal{M}(\mathbb{R}^n)$ be the collection of Borel probability measures on $\mathbb{R}^n$ with finite second moments.

II. PROBLEM FORMULATION

A. System Model

We consider discrete-time, linear time-varying systems evolving over a finite horizon $t = 0, \ldots, T - 1$, with dynamics

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \qquad (1)$$

where $x_t \in \mathbb{R}^n$ is the *system state*, $u_t \in \mathbb{R}^m$ is the *control input*, and $w_t \in \mathbb{R}^n$ is the *disturbance* acting on the system at time t. We assume perfect state feedback, and that the system matrices $A_t \in \mathbb{R}^{n \times n}$ and input matrices $B_t \in \mathbb{R}^{n \times m}$ for $t = 0, \ldots, T - 1$ are known at the outset. Both the initial system state x_0 and the disturbances $w_0, \ldots, w_{T-1}$ are assumed to be random variables whose joint distribution P is unknown but assumed to lie in a given compact set $\mathcal{P}$, termed the *ambiguity set*.

We denote the *state, input,* and *disturbance trajectories* by

$$
\begin{aligned}
x &:= (x_0, \ldots, x_T) \in \mathbb{R}^{N_x}, \\
u &:= (u_0, \ldots, u_{T-1}) \in \mathbb{R}^{N_u}, \\
w &:= (x_0, w_0, \ldots, w_{T-1}) \in \mathbb{R}^{N_x},
\end{aligned}
$$

where $N_x := n(T+1)$ and $N_u := mT$. To simplify notation, we have included the initial state x_0 as the first term in the disturbance trajectory w. With these definitions, the dynamics (1) can be expressed in terms of the following causal linear mapping from the input and disturbance trajectories to the state trajectory:

$$x = Fu + Gw, \qquad (2)$$

where $F \in \mathbb{R}^{N_x \times N_u}$ and $G \in \mathbb{R}^{N_x \times N_x}$ are block lower-triangular matrices. These matrices are straightforward to construct from the given state and input matrices (A_t, B_t), where $t = 0, \ldots, T - 1$.

The cost incurred by an input trajectory u and disturbance trajectory w is defined as

$$J(u, w) := x^\top Q x + u^\top R u, \qquad (3)$$

where $Q \in \mathbb{S}^{N_x}_+$ and $R \in \mathbb{S}^{N_u}_{++}$.

B. Linear Disturbance Feedback Controllers

In this paper, we consider the design of strictly causal linear disturbance feedback controllers[1] of the form

$$u_t = \sum_{k=0}^{t} K_{t,k} w_{k-1}, \quad \forall t = 0, \ldots, T - 1. \qquad (4)$$

Here, we have used the notational convention $w_{-1} := x_0$. The controller (4) can be expressed as a linear mapping from the disturbance trajectory to the input trajectory given by

$$u = Kw, \quad K \in \mathcal{K}, \qquad (5)$$

where $\mathcal{K} \subseteq \mathbb{R}^{N_u \times N_x}$ denotes the space of all block lower triangular matrices that correspond to the disturbance feedback parameterization specified in (4).

It is important to note that the family of strictly causal linear disturbance feedback controllers is equivalent to the family of causal linear state feedback controllers [20]. In particular, given a strictly causal disturbance feedback controller $K \in \mathcal{K}$, the change of variables[2] $L := (I + KG^{-1}F)^{-1}KG^{-1}$ induces an equivalent causal state feedback controller L that satisfies $Lx = Kw$ for all $w \in \mathbb{R}^{N_x}$ [22]. However, from an optimization standpoint, the disturbance feedback parameterization is preferred over the state feedback parameterization, because the cost function $J(Kw, w)$ resulting from the disturbance feedback parameterization is guaranteed to be convex with respect to the controller K [20].

C. Minimax Regret Optimal Control

We seek controllers belonging to the family $\mathcal{K}$ that minimize the worst-case expected regret across all probability distributions in the given ambiguity set $\mathcal{P}$. The *regret* incurred by a causal controller is defined as the difference between the cost it incurs and the cost incurred by the optimal noncausal controller (which knows the full disturbance trajectory at the outset). Formally, the regret incurred by a causal controller $K \in \mathcal{K}$ in response to a disturbance w is defined as

$$R(K, w) := J(Kw, w) - J(u^\star(w), w), \qquad (6)$$

---

[1]It is straightforward to extend the results of this paper to allow for the optimization over *affine* disturbance feedback policies. We forgo this more general treatment to streamline the presentation in this paper.

[2]This invertible nonlinear transformation is related to the well-known Youla parameterization in linear systems [21]. The matrix G is invertible since it is a lower triangular matrix with ones along its diagonal, and the matrix $I + KG^{-1}F$ is invertible since $KG^{-1}F$ is a block strictly lower triangular matrix.

where $u^\star : \mathbb{R}^{N_x} \to \mathbb{R}^{N_u}$ denotes the *optimal noncausal controller*, which is defined as

$$u^\star(w) := \underset{u \in \mathbb{R}^{N_u}}{\mathrm{argmin}}\, J(u, w). \tag{7}$$

With these definitions in hand, the *minimax regret optimal control* (MROC) problem can be formulated as

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P\left[R(K, w)\right]. \tag{8}$$

To solve the MROC problem, an explicit characterization of the optimal noncausal controller cost is needed. It is well known that the optimal noncausal controller exists, is unique, and is linear in the disturbance trajectory w (among all possible nonlinear controllers) [12], [23]. To see why this is true, note that $J(u, w)$ is a strictly convex quadratic function of u. It follows that the unique optimal noncausal controller can be obtained by solving the optimality condition $\nabla_u J(u, w) = 0$, which yields

$$u^\star(w) = K^\star w,$$

where $K^\star := -(R + F^\top Q F)^{-1} F^\top Q G$. Using the above identity, the cost incurred by the optimal noncausal controller can be expressed as

$$J(K^\star w, w) = w^\top G^\top (Q - QFD^{-1}F^\top Q)Gw,$$

where $D := R + F^\top Q F$ is a positive definite matrix. Furthermore, by substituting (2) into (3), and completing the square in (3) with respect to u, the cost function $J(u, w)$ can be rewritten as

$$J(u, w) = (u - K^\star w)^\top D(u - K^\star w) + J(K^\star w, w). \tag{9}$$

Notice that the first term in (9) represents the regret incurred by the control trajectory u in response to the disturbance w. Using the identity in (9), the MROC problem can be recast as

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P\left[w^\top (K - K^\star)^\top D(K - K^\star)w\right]. \tag{10}$$

From (10), it can be seen that the expected regret incurred by a controller $K \in \mathcal{K}$ under a distribution $P \in \mathcal{P}$ only depends on the distribution through its second moment.

D. Distributionally Robust Optimal Control

In Sections III and V, we compare the *minimax regret optimal* (MRO) controllers proposed in this paper with *distributionally robust optimal* (DRO) controllers, which are defined as solutions to the following optimization problem:

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P\left[J(Kw, w)\right]. \tag{11}$$

DRO controllers minimize the worst-case expected cost across all distributions in a given ambiguity set $\mathcal{P}$.

III. MOTIVATING EXAMPLE

Before delving into the main results of this paper, we first analyze a simple system that sheds light on the behavior of MRO and DRO controllers and the nature of the adversarial distributions they hedge against.

Method	Control Policy	Expected Cost
Optimal Noncausal	$-\left(\dfrac{1}{1+c}\right)(x_0 + w_0)$	$J^\star$
Optimal Causal	$-\left(\dfrac{1+\rho}{1+c}\right)x_0$	$J^\star + \dfrac{1-\rho^2}{1+c}$
MRO	$-\left(\dfrac{1}{1+c}\right)x_0$	$J^\star + \dfrac{1}{1+c}$
DRO	$-\left(\dfrac{2}{1+c}\right)x_0$	$J^\star + 2\left(\dfrac{1-\rho}{1+c}\right)$

TABLE I: Comparison of different control design methods applied to the specific system described in Section III for $c \geq 1$. The expected cost incurred by the optimal noncausal policy is given by $J^\star := 2c(1+\rho)/(1+c)$.

A. System Description

Consider a one-dimensional ($n = 1$) system that operates for a single time period ($T = 1$):

$$x_1 = x_0 + u_0 + w_0.$$

In this setting, there is a single control input u_0, and the disturbance trajectory $w = (x_0, w_0)$ is comprised of the initial state x_0 and initial disturbance w_0, which are assumed to have *known* marginal distributions with zero mean and unit variance. The correlation coefficient between the initial state and disturbance is assumed to be *unknown*. Given these assumptions, the ambiguity set of distributions can be expressed as

$$\mathcal{P} := \left\{ P \in \mathcal{M}(\mathbb{R}^2) \,\middle|\, \begin{array}{l} \mathbb{E}_P[w] = 0,\ \mathbb{E}_P[ww^\top] = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \\ \rho \in [-1, 1]\ \text{for}\ w \sim P \end{array} \right\},$$

where ρ denotes the correlation coefficient between x_0 and w_0. We consider a cost function that penalizes the terminal state and control input according to

$$J(u_0, w) = x_1^2 + cu_0^2,$$

where $c \geq 1$ is a given parameter.

We examine four different controllers: (i) the *optimal noncausal controller* defined in (7), (ii) the *MRO controller* defined in (8), (iii) the *DRO controller* defined in (11), and (iv) the *optimal causal controller*, which minimizes the expected cost given complete knowledge of the underlying distribution $P \in \mathcal{P}$. For the causal control design methods (ii)-(iv), we restrict our attention to linear policies of the form:

$$u_0 = -kx_0,$$

where $k \geq 0$ denotes the control gain. Closed-form expressions for the optimal controllers and their expected costs are provided in Table I. Due to space limitations, the derivation of these expressions is left as an exercise for the reader.
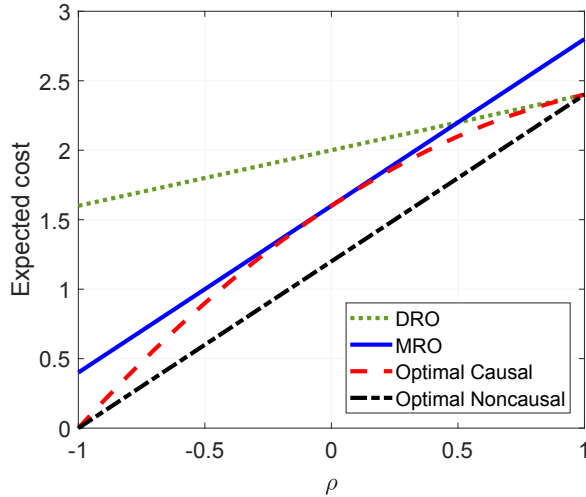
Fig. 1: Expected cost incurred by different controllers as a function of the correlation coefficient ρ for $c = 3/2$.

B. Discussion

In Fig. 1, we plot the expected cost of each controller as a function of the correlation coefficient ρ. The expected cost of the DRO controller is greatest at $\rho = 1$, i.e., when the initial state x_0 and the disturbance w_0 are perfectly positively correlated. This is unsurprising since heightened correlation between the initial state and the disturbance requires greater control effort to minimize the expected cost. Unlike the DRO controller, which minimizes the expected cost against the worst-case distribution in the ambiguity set, the MRO controller is found to strike a balance in compensating for both "favorable" distributions ($\rho = -1$) and "unfavorable" distributions ($\rho = 1$) in the ambiguity set, i.e., distributions that result in lower and higher expected costs, respectively. In fact, the MRO controller is equivalent to the optimal causal controller corresponding to the correlation coefficient $\rho = 0$. Thus, although the MRO controller incurs a greater expected cost than the DRO controller in the worst case, it strictly outperforms the DRO controller for all distributions in the ambiguity set with correlation coefficients $\rho \in [-1, 1/2)$. Moreover, the MRO controller achieves a constant expected regret of $1/(1 + c)$ over the ambiguity set, while the DRO controller exhibits a worst-case expected regret of $4/(1 + c)$ for distributions with $\rho = -1$.

IV. TRACTABLE CONVEX REFORMULATION

A primary challenge in solving the MROC problem (10) is that the inner maximization involves a supremum over an ambiguity set of distributions that may be nonconvex and infinite dimensional. To address this challenge, we consider ambiguity sets that are defined as the set of all distributions that are *close* to a nominal distribution with respect to the Wasserstein distance. The Wasserstein distance between two distributions can be interpreted as the minimum "work" or transportation cost required to move the probability mass from one distribution to the other. By defining the ambiguity set in this manner, we are able to draw on a well-

established duality theory for optimal transport problems to reformulate the inner maximization in (10) as a tractable finite-dimensional convex optimization problem [17], [24]. As one of our main contributions in Theorem 4, we provide an exact reformulation of the MROC problem (10) as a SDP. In Theorem 3, we also provide a structural characterization of a worst-case distribution achieving the supremum in the inner maximization in (10).

A. Strong Duality for Worst-Case Expectation Problems

Before presenting our main results, it is necessary to introduce some additional notation and define the type-2 Wasserstein distance between probability measures.

Definition 1 (Type-2 Wasserstein Distance). The *type-2 Wasserstein distance* between two distributions $P_1, P_2 \in \mathcal{M}(\mathbb{R}^{N_x})$ is defined as

$$W_2(P_1, P_2)^2 := \inf_{\pi \in \Pi(P_1, P_2)} \int_{\mathbb{R}^{N_x} \times \mathbb{R}^{N_x}} \|z_1 - z_2\|^2 \pi(dz_1, dz_2),$$

where $\Pi(P_1, P_2)$ denotes the set of all joint distributions in $\mathcal{M}(\mathbb{R}^{N_x} \times \mathbb{R}^{N_x})$ with marginal distributions P_1 and P_2.

Given a *nominal distribution* $P_0 \in \mathcal{M}(\mathbb{R}^{N_x})$, we define the ambiguity set $\mathcal{P}$ as the set of all distributions whose type-2 Wasserstein distance to P_0 is at most $r \geq 0$, i.e.,

$$\mathcal{P} := \{P \in \mathcal{M}(\mathbb{R}^{N_x}) \,|\, W_2(P, P_0) \leq r\}. \tag{12}$$

The *radius of the ambiguity set*, r, reflects the degree of uncertainty surrounding the accuracy of the nominal distribution. For example, when the central distribution is estimated from independent and identically distributed (i.i.d.) data, the radius of the ambiguity set can be selected to ensure that the ambiguity set contains the data-generating distribution with a desired confidence level, e.g., using the techniques described in [25]. If the radius of the ambiguity set is decreased to zero, then the ambiguity set collapses to a singleton set that only contains the nominal distribution, and the MROC problem (8) reduces to a conventional (ambiguity-free) stochastic optimal control problem given by $\inf_{K \in \mathcal{K}} \mathbb{E}_{P_0}[J(Kw, w)]$.

To solve the MROC problem (10), one must contend with the inner maximization (worst-case expectation) over the type-2 Wasserstein ball of distributions. Recently, it has been shown that strong duality holds for such infinite-dimensional maximization problems for a large family of objective functions [17], [24], [25]. In particular, the following known result from the literature provides conditions under which strong duality is guaranteed to hold, along with a useful characterization of the dual problem as a one-dimensional convex minimization problem.

Theorem 1 (Strong Duality, [17]). Let $r > 0$ and consider a family of worst-case expectation problems given by

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[f(w)],$$

where $f : \mathbb{R}^{N_x} \to \mathbb{R}$ is a Borel-measurable function such that $\mathbb{E}_{P_0}[|f(w)|] < \infty$. Then, strong duality holds, and

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[f(w)] = \inf_{\gamma \geq 0} \left\{ \gamma r^2 - \mathbb{E}_{P_0}[\phi(\gamma, w)] \right\}, \tag{13}$$

where the function $\phi : \mathbb{R} \times \mathbb{R}^{N_x} \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$\phi(\gamma, w) := \inf_{z \in \mathbb{R}^{N_x}} \{\gamma \|z - w\|^2 - f(z)\}.$$

Additionally, if $\inf\{\gamma \geq 0 \,|\, \mathbb{E}_{P_0}[\phi(\gamma, w)] > -\infty\} < \infty$, then the primal and dual optimal values in (13) are finite.

Building on Theorem 1, we provide conditions which enable the reformulation of worst-case expectation problems involving general quadratic objective functions as SDPs. We rely on the following assumption.

Assumption 1 (Nominal Distribution)**.** The nominal distribution P_0 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^{N_x}$.

Theorem 2 (Strong Duality for Quadratic Objectives)**.** Let $r > 0$ and Assumption 1 hold. Consider a family of worst-case expectation problems given by

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[w^\top C w], \tag{14}$$

where the matrix $C \in \mathbb{S}^{N_x}$ is assumed to satisfy $\lambda_{\max}(C) \neq 0$. Then, the optimal value of (14) is finite and equal to the optimal value of the following convex program:

$$\inf_{\gamma \geq 0} \{\gamma(r^2 - \mathrm{Tr}\,(M_0)) + \gamma^2 \,\mathrm{Tr}\,\left(M_0(\gamma I - C)^{-1}\right) \,|\, \gamma I \succ C\}, \tag{15}$$

where $M_0 := \mathbb{E}_{P_0}[ww^\top]$.

Proof. For any matrix $C \in \mathbb{S}^{N_x}$, the function $z \mapsto z^\top C z$ is Borel measurable, and satisfies $\mathbb{E}_{P_0}[|w^\top C w|] < \infty$ since the nominal distribution P_0 has a finite second moment by assumption. Then, by Theorem 1, it holds that

$$\sup_{P \in \mathcal{P}} \mathbb{E}_P[w^\top C w] = \inf_{\gamma \geq 0} \left\{\gamma r^2 - \mathbb{E}_{P_0}\left[\phi(\gamma, w)\right]\right\}, \tag{16}$$

where

$$\phi(\gamma, w) = \inf_{z \in \mathbb{R}^{N_x}} \left\{z^\top(\gamma I - C)z - 2\gamma w^\top z\right\} + \gamma \|w\|^2.$$

Also, by Theorem 1, the primal and dual optimal values in (16) are finite if $\inf\{\gamma \geq 0 \,|\, \mathbb{E}_{P_0}[\phi(\gamma, w)] > -\infty\} < \infty$. Under the conditions stated in Theorem 2, we shall prove that $\mathbb{E}_{P_0}[\phi(\gamma, w)] > -\infty$ if and only if $\gamma I - C \succ 0$. The implication in the "if" direction is clear.

We prove the implication in the "only if" direction using contraposition, i.e., we shall prove that $\mathbb{E}_{P_0}[\phi(\gamma, w)] = -\infty$ if $\gamma I - C \not\succ 0$. To streamline notation, let $C_\gamma := \gamma I - C$. Additionally, let $v \in \mathbb{R}^{N_x}$ denote an eigenvector of the matrix C_γ associated with the smallest eigenvalue $\lambda_{\min}(C_\gamma)$.

Note that $C_\gamma \not\succ 0$ is equivalent to $\lambda_{\min}(C_\gamma) \leq 0$. If $\lambda_{\min}(C_\gamma) < 0$, then $\inf_{z \in \mathbb{R}^{N_x}} \left\{z^\top C_\gamma z - 2\gamma w^\top z\right\} = -\infty$ for any $w \in \mathbb{R}^{N_x}$, which implies that $\mathbb{E}_{P_0}[\phi(\gamma, w)] = -\infty$. If $\lambda_{\min}(C_\gamma) = 0$, then

$$\inf_{z \in \mathbb{R}^{N_x}} \left\{z^\top C_\gamma z - 2\gamma w^\top z\right\}$$
$$\leq \inf_{\theta \in \mathbb{R}} \left\{(\theta v)^\top C_\gamma (\theta v) - 2\gamma w^\top (\theta v)\right\}$$
$$= \inf_{\theta \in \mathbb{R}} \left\{-2\gamma w^\top (\theta v)\right\}$$

$$= \inf_{\theta \in \mathbb{R}} \left\{-2\lambda_{\max}(C) w^\top (\theta v)\right\}$$
$$= -\infty \; P_0\text{-almost surely}.$$

The first equality follows from the assumption that v is an eigenvector of C_γ associated with the eigenvalue $\lambda_{\min}(C_\gamma) = 0$. The second equality follows from $\lambda_{\min}(C_\gamma) = 0$, which is equivalent to $\gamma = \lambda_{\max}(C)$. The final equality follows from the assumption that $\lambda_{\max}(C) \neq 0$ and Assumption 1 (which implies that $w^\top v \neq 0$ P_0-almost surely). This proves that $\mathbb{E}_{P_0}[\phi(\gamma, w)] > -\infty$ if and only if $C_\gamma \succ 0$. It follows that

$$\inf\{\gamma \geq 0 \,|\, \mathbb{E}_{P_0}[\phi(\gamma, w)] > -\infty\} = \inf\{\gamma \geq 0 \,|\, C_\gamma \succ 0\}$$
$$= \max\{\lambda_{\max}(C), 0\}$$
$$< \infty,$$

proving finiteness of the primal and dual optimal values in (16).

We complete the proof by showing that the right-hand side of (16) is equal to (15). From previous arguments, it follows that

$$\inf_{\gamma \geq 0} \left\{\gamma r^2 - \mathbb{E}_{P_0}\left[\phi(\gamma, w)\right]\right\}$$
$$= \inf_{\gamma \geq 0} \left\{\gamma r^2 - \mathbb{E}_{P_0}\left[\phi(\gamma, w)\right] | C_\gamma \succ 0\right\},$$

which leads to (15) as $\inf_{z \in \mathbb{R}^{N_x}} \left\{z^\top C_\gamma z - 2\gamma w^\top z\right\} = -\gamma^2 w^\top C_\gamma^{-1} w$ under the condition $C_\gamma \succ 0$. $\qquad\square$

Together with Assumption 1, the condition $\lambda_{\max}(C) \neq 0$ ensures that the dual optimal value (15) is finite. While this condition permits a variety of quadratic objective functions, including definite and indefinite quadratic functions, it does exclude negative semidefinite quadratic functions with $\lambda_{\max}(C) = 0$.

We also note that Theorem 2 extends a number of related results in the literature [18], [19] by expanding the family of nominal distributions under which the worst-case expectation problem (14) is guaranteed to admit an equivalent convex reformulation as (15). In particular, the convex reformulation provided in [19, Theorem 16] relies on the assumption that the nominal distribution is a (possibly degenerate) elliptical distribution, while Assumption 1 only requires that the nominal distribution be absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^{N_x}$. While the family of probability distributions satisfying Assumption 1 is quite broad, it does rule out nominal distributions supported on lower dimensional manifolds (e.g., degenerate distributions), and distributions supported on finite sets. For a comprehensive treatment of nominal distributions supported on finite sets, we refer the reader to [25], which provides tractable convex reformulations for an array of worst-case expectation problems over type-1 Wasserstein balls centered at empirical distributions.

B. Worst-Case Distribution Characterization

Using the dual reformulation of the quadratic worst-case expectation problem (14) provided in Theorem 2, we

now characterize a worst-case distribution which attains the optimal value of the worst-case expectation problem (14).

Theorem 3 (Worst-Case Distribution). Let the assumptions of Theorem 2 hold and let $\gamma^\star \geq 0$ satisfy $\gamma^\star I \succ C$ and be a solution to the algebraic equation

$$\text{Tr}\big((\gamma(\gamma I - C)^{-1} - I)^2 M_0\big) = r^2. \qquad (17)$$

Let $P^\star \in \mathcal{M}(\mathbb{R}^{N_x})$ denote the distribution of

$$w^\star := \gamma^\star(\gamma^\star I - C)^{-1}w, \qquad (18)$$

where $w \sim P_0$. Then, the distribution $P^\star$ belongs to the ambiguity set $\mathcal{P}$ and it attains the optimal value of the worst-case expectation problem (14).

Proof. To prove the desired result, we need to show that $P^\star$ is both feasible and optimal in (14). To streamline notation, let $C_\gamma := \gamma I - C$.

First, we show that the distribution $P^\star$ belongs to the ambiguity set $\mathcal{P}$ by verifying that it satisfies $W_2(P^\star, P_0) \leq r$. To achieve this, let $\pi \in \Pi(P^\star, P_0)$ be the joint distribution of $w^\star$ and w. Then, we have that

$$W_2(P^\star, P_0)^2$$
$$\leq \mathbb{E}_\pi\big[(w^\star - w)^\top(w^\star - w)\big]$$
$$= \text{Tr}\big(\mathbb{E}_{P^\star}[w^\star w^{\star\top}] - 2\mathbb{E}_\pi[ww^{\star\top}] + \mathbb{E}_{P_0}[ww^\top]\big)$$
$$\overset{(a)}{=} \text{Tr}\big(\gamma^{\star2}C_{\gamma^\star}^{-1}M_0 C_{\gamma^\star}^{-1} - 2\gamma^\star C_{\gamma^\star}^{-1}M_0 + M_0\big)$$
$$= \text{Tr}\big((\gamma^\star C_{\gamma^\star}^{-1} - I)^2 M_0\big)$$
$$\overset{(b)}{=} r^2,$$

where (a) holds by (18) and (b) holds by (17). Hence, $P^\star \in \mathcal{P}$ and is feasible for (14).

Next, we prove that the distribution $P^\star$ is optimal for the worst-case expectation problem (14). Let $g : \mathbb{R}_+ \to \mathbb{R}$ denote the dual objective function in (15). It can be expressed as

$$g(\gamma) = \gamma(r^2 - \text{Tr}(M_0)) + \gamma^2 \text{Tr}(M_0 C_\gamma^{-1}).$$

The function g is strictly convex on the interval $(\lambda_{\max}(C), \infty)$ since its second derivative satisfies

$$\frac{d^2 g(\gamma)}{d\gamma^2} = 2\,\text{Tr}(M_0 C^2 C_\gamma^{-3}) > 0$$

for all $\gamma \in (\lambda_{\max}(C), \infty)$. Also, note that g goes to infinity as γ tends to infinity or approaches $\lambda_{\max}(C)$ from the right. Consequently, the dual function g has a unique minimizer $\gamma^\star$ in the interval $(\lambda_{\max}(C), \infty)$. Hence, the first order optimality condition is a necessary and sufficient condition for optimality on the interval $(\lambda_{\max}(C), \infty)$. This condition is given by

$$\frac{dg(\gamma)}{d\gamma} = r^2 - \text{Tr}\big(M_0 - 2\gamma C_\gamma^{-1}M_0 + \gamma^2 C_\gamma^{-2}M_0\big) = 0,$$

which can be rewritten as (17).

By Theorem 2, the optimal value of the primal problem (14) is finite and equal to the optimal value of the dual problem (15). Hence, to prove that $P^\star$ attains the supremum

of (14), it suffices to show that $\mathbb{E}_{P^\star}[w^{\star\top}Cw^\star] = g(\gamma^\star)$. It holds that

$$g(\gamma^\star) \overset{(a)}{=} \gamma^\star \text{Tr}\big((\gamma^\star C_{\gamma^\star}^{-1} - I)^2 M_0 - M_0 + \gamma^\star C_{\gamma^\star}^{-1}M_0\big)$$
$$= \gamma^\star \text{Tr}\big((-\gamma^\star C_{\gamma^\star}^{-1} + \gamma^{\star2}C_{\gamma^\star}^{-2})M_0\big)$$
$$= \gamma^{\star2}\text{Tr}\big(C_{\gamma^\star}^{-1}M_0(-I + \gamma^\star C_{\gamma^\star}^{-1})\big)$$
$$\overset{(b)}{=} \text{Tr}\big(\gamma^{\star2}C_{\gamma^\star}^{-1}M_0 C_{\gamma^\star}^{-1}C\big)$$
$$\overset{(c)}{=} \mathbb{E}_{P^\star}[w^{\star\top}Cw^\star],$$

where (a) follows from plugging (17) into $g(\gamma^\star)$; (b) holds since $-I + \gamma^\star C_{\gamma^\star}^{-1} = -C_{\gamma^\star}^{-1}C_{\gamma^\star} + \gamma^\star C_{\gamma^\star}^{-1} = C_{\gamma^\star}^{-1}C$; and (c) follows from (18). Therefore, $P^\star$ is optimal and the desired result holds. □

In Theorem 3, the proposed worst-case distribution $P^\star$ is given by the push-forward measure of P_0 through the linear transformation $z \mapsto \gamma^\star(\gamma^\star I - C)^{-1}z$. It is also possible to show that $P^\star$ is an extremal distribution in the ambiguity set $\mathcal{P}$. In particular, by using a well-known lower bound on the type-2 Wasserstein distance due to Gelbrich [26, Theorem 2.1], one can show that $W_2(P^\star, P_0) \geq r$.

C. Semidefinite Programming Reformulation

Building on Theorem 2, we now show how to reformulate the MROC problem (8) as a SDP in Theorem 4.

Theorem 4 (SDP Reformulation of MROC Problem). Let $r > 0$ and suppose that $K^\star \notin \mathcal{K}$ and Assumption 1 holds. Then, the MROC problem (8) can be equivalently reformulated as the following SDP:

$$\inf \ \gamma(r^2 - \text{Tr}(M_0)) + \text{Tr}(X) \qquad (19a)$$
$$\text{s.t. } K \in \mathcal{K}, \gamma \geq 0, X \in \mathbb{S}_+^{N_x},$$
$$\begin{bmatrix} \gamma I & (K - K^\star)^\top \\ K - K^\star & D^{-1} \end{bmatrix} \succ 0, \qquad (19b)$$
$$\begin{bmatrix} X & \gamma M_0^{\frac{1}{2}} & 0_{N_u \times N_x}^\top \\ \gamma M_0^{\frac{1}{2}} & \gamma I & (K - K^\star)^\top \\ 0_{N_u \times N_x} & K - K^\star & D^{-1} \end{bmatrix} \succeq 0, \qquad (19c)$$

where $0_{N_u \times N_x}$ denotes a N_u-by-N_x matrix of all zeros and the decision variables are K, γ and X.

Proof. To streamline notation, define the matrix

$$C_K := (K - K^\star)^\top D(K - K^\star).$$

The condition $K^\star \notin \mathcal{K}$ ensures that $\lambda_{\max}(C_K) > 0$ for all $K \in \mathcal{K}$. This, together with Assumption 1, allows us to apply Theorem 2 to reformulate the MROC problem as

$$\inf_{K \in \mathcal{K}} \sup_{P \in \mathcal{P}} \mathbb{E}_P[w^\top C_K w]$$
$$= \inf_{K \in \mathcal{K}, \gamma \geq 0} \big\{\gamma(r^2 - \text{Tr}(M_0))$$
$$+ \gamma^2 \text{Tr}(M_0(\gamma I - C_K)^{-1}) \mid \gamma I \succ C_K\big\}. \quad (20)$$

Since $D \succ 0$, it follows from the Schur complement theorem that the constraint $\gamma I \succ C_K$ is equivalent to the linear matrix inequality (19b). Now, observe that the second term

$\gamma^2 \operatorname{Tr}\left(M_0(\gamma I - C_K)^{-1}\right)$ in the right-hand side of (20) can be expressed as

$$\min_{X \succeq 0}\left\{ \operatorname{Tr}(X) \mid X \succeq \gamma M_0^{\frac{1}{2}}(\gamma I - C_K)^{-1}\gamma M_0^{\frac{1}{2}}\right\}.$$

Once again, using the Schur complement theorem, we can reformulate the constraint $X \succeq \gamma M_0^{\frac{1}{2}}(\gamma I - C_K)^{-1}\gamma M_0^{\frac{1}{2}}$ as

$$\begin{bmatrix} X & \gamma M_0^{\frac{1}{2}} \\ \gamma M_0^{\frac{1}{2}} & \gamma I - C_K \end{bmatrix} \succeq 0, \qquad (21)$$

since $\gamma I - C_K \succ 0$. To prove equivalence between the matrix inequality (21) and (19c), it is helpful to rewrite (21) as

$$\begin{bmatrix} X & \gamma M_0^{\frac{1}{2}} \\ \gamma M_0^{\frac{1}{2}} & \gamma I \end{bmatrix} - \begin{bmatrix} 0_{N_u \times N_x}^\top \\ (K - K^\star)^\top \end{bmatrix} D \begin{bmatrix} 0_{N_u \times N_x}^\top \\ (K - K^\star)^\top \end{bmatrix}^\top \succeq 0.$$

Since $D \succ 0$, it follows from the Schur complement theorem that the above constraint is equivalent to (19c). This proves that problem (20) is equivalent to the SDP (19) and completes the proof. $\qquad\square$

Note that the assumption $K^\star \notin \mathcal{K}$ is without loss of generality. If the optimal noncausal controller $K^\star$ happens to be causal, i.e., $K^\star \in \mathcal{K}$, then the minimax regret optimal controller is just $u = K^\star w$, which avoids the need to solve problem (8).

The size of the SDP in (19) grows polynomially with the state dimension n, input dimension m, and time horizon T. Although SDPs can be solved with arbitrary accuracy in polynomial time using interior point methods, in practice they are expensive to solve in high dimensions. To improve scalability, one can exploit sparsity in the problem data [27], [28], or employ approximations of the positive semidefinite cone [29], [30].

V. NUMERICAL EXPERIMENTS

In this section, we compare the MRO and DRO control design methods within the context of data-driven control. We consider a one-dimensional ($n = 1$) controlled random walk with dynamics given by

$$x_{t+1} = x_t + u_t + w_t$$

for $t = 0, ..., T - 1$. In our experiments, we consider a time horizon of $T = 10$ and take the cost matrices to be $Q, R = I$. The disturbance trajectory w is assumed to have a Gaussian distribution given by $\mathcal{N}(\mu, I)$.

We assume that the controllers have access to N i.i.d. samples $w^{(1)}, \ldots, w^{(N)}$ of the disturbance trajectory, but do not know the underlying distribution from which the data is generated. Using these data, we consider an ambiguity set $\mathcal{P}$ that is centered at a nominal distribution $P_0 \in \mathcal{M}(\mathbb{R}^{N_x})$ whose second moment matrix is given by the empirical estimate:

$$M_0 = \frac{1}{N}\sum_{i=1}^{N} w^{(i)}w^{(i)\top}. \qquad (22)$$

Note that the SDP reformulations of the MROC and DROC problems only depend on the second moment matrix of the
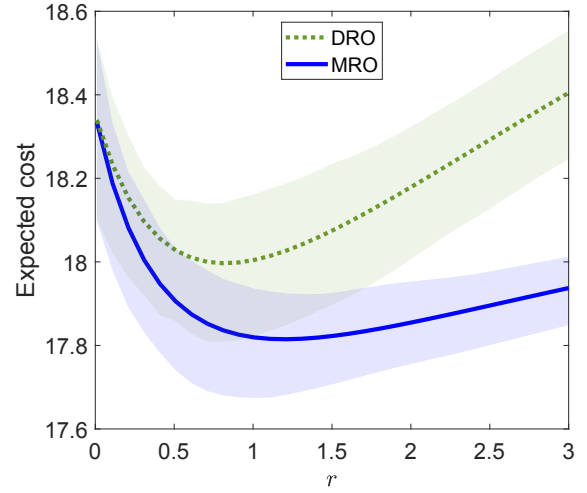
(a) $\mu = 0$
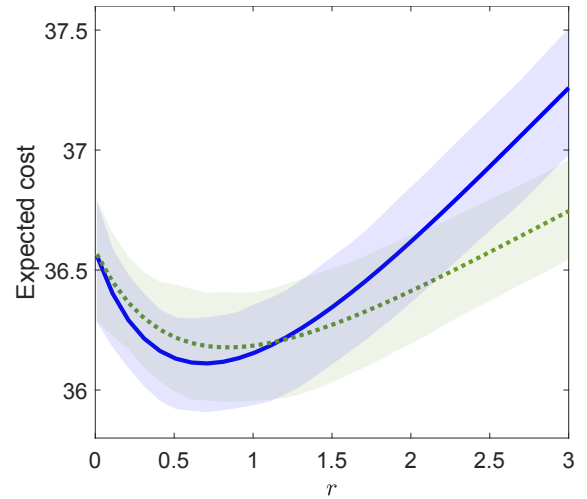
(b) $\mu = 1$

Fig. 2: These plots depict the expected cost (averaged over 100 trials) incurred by the MRO and DRO controllers under the true data-generating distribution $\mathcal{N}(\mu, I)$ versus the ambiguity set radius r. The shaded regions depict the range between the corresponding 20th and 80th empirical quantiles.

nominal distribution, so only specifying its second moment suffices. In our experiments, we vary the radius r of the ambiguity set to investigate its effect on the performance of the resulting MRO and DRO controllers.

We examine two different data-generating distributions: one that has a zero mean vector ($\mu = 0$) and one that has a constant nonzero mean vector of all ones ($\mu = 1$). For each distribution, we conduct 100 independent trials. In each trial, we draw $N = 50$ i.i.d. samples of the disturbance trajectory from the underlying distribution $\mathcal{N}(\mu, I)$. We vary the radius of the ambiguity set from zero to three and compute the MRO controller for each radius value by solving the SDP (19). We compute the DRO controller by solving a SDP reformulation of (11) that follows from Theorem 4.

In Fig. 2, we plot the expected costs (averaged over

100 trials) incurred by the MRO and DRO controllers as a function of the ambiguity set radius. In each case, the expectations are calculated with respect to the true data-generating distribution.

When the radius r of the ambiguity set is equal to zero, the MRO and DRO control problems are identical to the ambiguity-free stochastic control problem given by $\inf_{K \in \mathcal{K}} \mathbb{E}_{P_0}[J(Kw, w)]$. This can be interpreted as a *certainty equivalent* approach to data-driven control. For small values of r, we observe that both the MRO and DRO controllers improve upon the performance of the certainty equivalent controller. In the zero-mean case (Fig. 2a), the MRO controller strictly outperforms the DRO controller for all radii in the specified range. In the nonzero-mean case (Fig. 2b), the MRO controller outperforms the DRO controller for small radius values, while the DRO controller outperforms the MRO controller for larger radii. This reveals that, in some settings, the choice of ambiguity set radius r can have an important effect on the superiority of one method over the other. That said, for both distributions, the MRO controller achieves the smallest expected cost over the specified radius range, compared to the DRO controller.

VI. Conclusion

In this paper, we propose a distributionally robust approach to synthesizing minimax regret optimal controllers for linear discrete-time systems subject to stochastic additive disturbances whose probability distribution is only known to lie in a type-2 Wasserstein ball of distributions. Building on existing strong duality results for Wasserstein distributionally robust optimization problems, we provide an equivalent reformulation of the minimax regret optimal control problem as a tractable semidefinite program. As a direction for future research, it would be interesting to generalize the framework proposed in this paper to enable the design of minimax regret optimal controllers with partial state feedback.

References

[1] L. J. Savage, "The theory of statistical decision," *Journal of the American Statistical association*, vol. 46, no. 253, pp. 55–67, 1951.

[2] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1931–1946, 2004.

[3] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation of deterministic parameters with bounded data uncertainties," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2177–2188, 2004.

[4] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends® in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[5] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends® in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2012.

[6] Y. Abbasi-Yadkori and C. Szepesvári, "Regret bounds for the adaptive control of linear quadratic systems," in *Proceedings of the 24th Annual Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, vol. 19, 2011, pp. 1–26.

[7] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "Regret bounds for robust adaptive control of the linear quadratic regulator," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[8] Y. Ouyang, M. Gagrani, and R. Jain, "Posterior sampling-based reinforcement learning for control of unknown linear systems," *IEEE Trans. on Automatic Control*, vol. 65, no. 8, pp. 3600–3607, 2019.

[9] O. Sabag, G. Goel, S. Lale, and B. Hassibi, "Regret-optimal controller for the full-information problem," in *2021 American Control Conference (ACC)*, 2021, pp. 4777–4782.

[10] G. Goel and B. Hassibi, "Regret-optimal estimation and control," *IEEE Trans. on Automatic Control*, vol. 68, no. 5, pp. 3041–3053, 2023.

[11] A. Didier, J. Sieber, and M. N. Zeilinger, "A system level approach to regret optimal control," *IEEE Control Systems Letters*, vol. 6, pp. 2792–2797, 2022.

[12] A. Martin, L. Furieri, F. Dörfler, J. Lygeros, and G. Ferrari-Trecate, "Safe control with minimal regret," in *Proceedings of the 4th Annual Learning for Dynamics and Control Conference*, ser. Proceedings of Machine Learning Research, vol. 168, 2022, pp. 726–738.

[13] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.

[14] B. P. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 430–442, 2015.

[15] P. Coppens, M. Schuurmans, and P. Patrinos, "Data-driven distributionally robust LQR with multiplicative noise," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, ser. Proceedings of Machine Learning Research, vol. 120, 2020, pp. 521–530.

[16] K. Kim and I. Yang, "Distributional robustness in minimax linear quadratic control with Wasserstein distance," *SIAM Journal on Control and Optimization*, vol. 61, no. 2, pp. 458–483, 2023.

[17] R. Gao and A. Kleywegt, "Distributionally robust stochastic optimization with Wasserstein distance," *Mathematics of Operations Research*, vol. 48, no. 2, pp. 603–655, 2023.

[18] V. A. Nguyen, D. Kuhn, and P. Mohajerin Esfahani, "Distributionally robust inverse covariance estimation: The Wasserstein shrinkage estimator," *Operations research*, vol. 70, no. 1, pp. 490–515, 2022.

[19] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, *Wasserstein Distributionally Robust Optimization: Theory and Applications in Machine Learning*. Informs, 2019, pp. 130–166.

[20] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

[21] D. Youla, H. Jabr, and J. Bongiorno, "Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case," *IEEE Transactions on Automatic Control*, vol. 21, no. 3, pp. 319–338, 1976.

[22] J. Skaf and S. P. Boyd, "Design of affine controllers via convex optimization," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2476–2487, 2010.

[23] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic estimation and control: a unified approach to H_2 and H_∞ theories*. SIAM, 1999.

[24] J. Blanchet and K. Murthy, "Quantifying distributional model risk via optimal transport," *Mathematics of Operations Research*, vol. 44, no. 2, pp. 565–600, 2019.

[25] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

[26] M. Gelbrich, "On a formula for the L^2 Wasserstein metric between measures on Euclidean and Hilbert spaces," *Mathematische Nachrichten*, vol. 147, no. 1, pp. 185–203, 1990.

[27] L. Vandenberghe, M. S. Andersen *et al.*, "Chordal graphs and semidefinite optimization," *Foundations and Trends® in Optimization*, vol. 1, no. 4, pp. 241–433, 2015.

[28] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn, "Chordal decomposition in operator-splitting methods for sparse semidefinite programs," *Mathematical Programming*, vol. 180, no. 1-2, pp. 489–532, 2020.

[29] R. Louca and E. Bitar, "A hierarchy of polyhedral approximations of robust semidefinite programs," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 7056–7062.

[30] A. A. Ahmadi and A. Majumdar, "DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization," *SIAM Journal on Applied Algebra and Geometry*, vol. 3, no. 2, pp. 193–230, 2019.