2024 American Control Conference (ACC)
July 8-12, 2024. Toronto, Canada

Optimal Design of Disturbance Attenuation Feedback Controllers for Linear Dynamical Systems*

Davide Mannini[1], Robin Strässer[2], and James B. Rawlings[1]

Abstract— **This paper finds the optimal feedback controller for the discrete time, finite horizon disturbance attenuation problem under bounded disturbances. We consider a linear dynamical system and a quadratic objective function with a resulting nonlinear optimization and optimal nonlinear controller. In the space of initial states, two regions are identified. One region, containing the zero initial state, features the linear optimal H_∞ controller, while the other region features nonlinear optimal control, and converges to the linear quadratic regulator (LQR) controller for large initial states. The transition between the two regions offers a unified framework that spans from H_∞ control to LQR control as a function of the relative magnitude of the initial state and the imposed disturbance. This study enhances the versatility of disturbance attenuation feedback controllers and expands on the previous work on the dynamic game theory approach to optimal robust control design.**

I. INTRODUCTION

The disturbance attenuation problem objective is the design of a feedback control law that guarantees the effect of a disturbance on a dynamical system is reduced to an acceptable level. Disturbance attenuation has been considered to be of practical importance in several engineering and scientific fields, from engine control in mechanical engineering to distillation column control in chemical engineering [1]. Successful industrial applications of robust controls using disturbance attenuation include helicopter control design [2].

Zames [3] introduced in the frequency domain the formulation of a disturbance attenuation problem for linear dynamical systems as an H_∞ optimization problem, which minimizes the H_∞ norm of the transfer function between the bounded disturbance input and the performance output. Limitations of frequency domain solution methods motivated Glover and Doyle [4] to develop the time domain interpretation of the H_∞ optimization problem. Basar [5] further expanded time domain disturbance attenuation optimization with a game theoretic approach.

Time domain solution methods typically consider zero initial states or a small neighborhood around the zero initial state. Didinsky and Basar [6] considered the design of a feedback controller for a dynamical linear system with nonzero initial states, proposing two solution regions in the space of initial states. Basar [7] also suggested that the optimal control is nonlinear in a neighborhood around the zero initial state. Finite and infinite horizon recursive and non-recursive methods for the optimal design of disturbance attenuation feedback controllers are established. However, efficient solution algorithms covering the entire range of relevant initial states and bounded disturbances are currently lacking.

In this article, we expand upon the previous disturbance attenuation and H_∞ control work and introduce a single, scalar decision variable minimization to directly solve the disturbance attenuation problem under bounded disturbances for any initial state. The resulting optimal feedback controller is *nonlinear*. Two regions in the space of initial states are defined: one, containing the zero initial state, where the optimal control is linear and equivalent to the optimal H_∞ control, and one where the optimal control is nonlinear, and converges to the optimal LQR control as the initial state becomes large.

Notation: Let $\mathbb{I}$ and $\mathbb{R}$ denote the integers and reals. The 2-norm of a vector x is defined as $|x| = \sqrt{x^* x}$, and the norm of a signal, defined as vector-valued function of scalar-valued time t, $x(t)$, is defined as $\|x(t)\|_2 := (\sum_{k=0}^{\infty} |x^*(k)x(k)|)^{1/2}$. Let $\mathbf{x} \in \mathbb{X}^N$ be a vector sequence defined as $\mathbf{x} := (x(0), x(1), \dots, x(N-1))$.

II. DISTURBANCE ATTENUATION PROBLEM

The original disturbance attenuation problem in the frequency domain attempted to find the controller that minimizes the H_∞ norm of the transfer function, $T(s)$, from a disturbance input, $\bar{w}(s)$, to a performance output, $\bar{z}(s)$ [3]. The H_∞ norm of the transfer function is defined as

$$\|T(s)\|_\infty = \sup_{\bar{w}(s)\neq 0} \frac{\|\bar{z}(s)\|_2}{\|\bar{w}(s)\|_2} \tag{1}$$

We find a time domain interpretation of definition (1) by applying Plancherel's theorem [8, p.69] and squaring the norms of the performance output and disturbance input

$$\sup_{\bar{w}(s)\neq 0} \frac{\|\bar{z}(s)\|_2^2}{\|\bar{w}(s)\|_2^2} = \sup_{w(t)\neq 0} \frac{\|z(t)\|_2^2}{\|w(t)\|_2^2} \tag{2}$$

We constrain the squared norm of the disturbance input in (2) and the time domain disturbance attenuation problem relevant for this work is

$$\inf_{u(t)} \sup_{\|w(t)\|_2^2 \leq \alpha} \frac{\|z(t)\|_2^2}{\|w(t)\|_2^2} \tag{3}$$

Where $u(t)$ is the control input and α is a scalar parameter. Based on the structure of problem (3), we define a general

*The authors gratefully acknowledge the financial support of the National Science Foundation (NSF) under Grant No. 2027091.

[1] Department of Chemical Engineering, University of California, Santa Barbara, 3357 Engrg II UCSB, Santa Barbara, CA 93106, USA dmannini@ucsb.edu, jbraw@ucsb.edu

[2] Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany robin.straesser@ist.uni-stuttgart.de

979-8-3503-8265-5/$31.00 ©2024 AACC

disturbance attenuation for the following discrete time system

$$x^+ = f(x,u,w) \qquad f : \mathbb{R}^n \times \mathbb{U} \times \mathbb{W} \to \mathbb{R}^n$$

in which $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{U} \subseteq \mathbb{R}^m$ is the controlled input, $w \in \mathbb{W} \subseteq \mathbb{R}^q$ is the disturbance, and x^+ denotes the successor state. The finite horizon optimal control problem is

$$\mathscr{P}_N(x): \qquad V_N^0(x) = \inf_{\mathbf{u}\in\mathscr{U}(x)} \sup_{\mathbf{w}\in\mathbb{W}^N} V_N(x,\mu,\mathbf{w})$$

where

$$V_N(x,\mu,\mathbf{w}) := \sum_{i=0}^{N-1} \ell(x(i),u(i)) + V_f(x(N))$$

$$\mathscr{U}(x) := \{\mathbf{u} \in \mathbb{U} \mid (x(i),u(i)) \in \mathbb{R}^n \times \mathbb{R}^m, \; \forall i \in \mathbb{I}_{0:N-1}, \; x(N) \in \mathbb{X}_f, \; \forall \mathbf{w} \in \mathbb{W}^N\}$$

For the horizon $N \in \mathbb{I}_{\geq 0}$, we use $x(i) := \phi(i;x,\mu,\mathbf{w})$ to denote the state trajectory at time $i \in \mathbb{I}_{0:N}$, given the initial state $x(0) = x \in \mathbb{R}^n$, and $u(i) = \mu(x(i))$ to denote the control input trajectory at time $i \in \mathbb{I}_{0:N-1}$. Let $\mathbf{u}^0(x)$ denote the minimizing value of the control variable $\mathbf{u}$, $\mu^0(x)$ the corresponding optimal control policy, and let $V_N^0(x) := V_N(x,\mu^0(x))$ denote the value function. We implicitly assume that a solution to $\mathscr{P}_N(x)$ exists for all $x \in \mathscr{X}_N(x) := \{x \mid \mathscr{U}(x) \neq \emptyset\}$ and that $\mathscr{X}_N$ is not empty. We thus replace the operators inf and sup with min and max, respectively. In this paper we will be dealing with a structure analogous to (3) and performance output

$$z(t) = \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$$

with an unconstrained linear dynamical system and quadratic objective function

$$f(x,u,w) = Ax + Bu + w \qquad x(0) = x_0$$

$$|z|^2 = \ell(x,u) = \frac{1}{2}(x'Qx + u'Ru) \qquad V_f(x) = \frac{1}{2}x'P_f x$$

in which $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $w \in \mathbb{W}$, $Q > 0$, $R > 0$, $P_f > 0$, and x_0 is the initial state. We parameterize the control input in feedback form

$$u = \mu(x,v) = \bar{K}x + v$$

where $\bar{K}$ is an arbitrary control gain and $v \in \mathbb{R}^m$ is a control variable. The optimal control problem reduces to

$$V_N(x,\mathbf{v},\mathbf{w}) = \sum_{i=0}^{N-1} \ell(x(i),u(i)) + V_f(x(N))$$

$$\mathscr{V}_N(x) := \{\mathbf{v} \mid (x(i),u(i)) \in \mathbb{R}^n \times \mathbb{R}^m, \; x(N) \in \mathbb{X}_f, \; \forall \mathbf{w} \in \mathbb{W}^N\}$$

Where the disturbance set $\mathbb{W}$ is defined as

$$\mathbb{W}^0 := \{\mathbf{w} \mid \|\mathbf{w}\|_2^2 \leq \alpha\}$$

Where the scalar α is the bound disturbance. From the linearity of the system, the optimal solution of the maximization over $\mathbf{w}$ lies on the boundary of the constraint set. So we can equivalently express the constraint as

$$\mathbb{W}^0 := \{\mathbf{w} \mid \|\mathbf{w}\|_2^2 = \alpha\}$$

Solving the model as a finite horizon non-recursive problem

$$\mathbf{x} = \mathscr{A}x_0 + \mathscr{B}\mathbf{v} + \mathscr{G}\mathbf{w}$$

with

$$\mathbf{x} := \begin{bmatrix} x(1) \\ x(2) \\ x(3) \\ \vdots \\ x(N) \end{bmatrix} \quad \mathbf{v} := \begin{bmatrix} v(0) \\ v(1) \\ v(2) \\ \vdots \\ v(N-1) \end{bmatrix} \quad \mathbf{w} := \begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ \vdots \\ w(N-1) \end{bmatrix}$$

$$\mathscr{A} := \begin{bmatrix} A_K \\ A_K^2 \\ A_K^3 \\ \vdots \\ A_K^N \end{bmatrix} \quad \mathscr{B} := \begin{bmatrix} B & 0 & 0 & \cdots & 0 \\ A_K B & B & 0 & \cdots & 0 \\ A_K^2 B & A_K B & B & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_K^{N-1}B & A_K^{N-2}B & A_K^{N-3} & \cdots & B \end{bmatrix}$$

$$\mathscr{G} := \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ A_K & I & 0 & \cdots & 0 \\ A_K^2 & A_K & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_K^{N-1} & A_K^{N-2} & A_K^{N-3} & \cdots & I \end{bmatrix}$$

where $A_K = A + B\bar{K}$. Next, we evaluate the control $\mathbf{u}$ given the control parameterization $u = \bar{K}x + v$

$$\mathbf{u} = \mathscr{K}_1 x_0 + \mathscr{K}_2\mathbf{x} + \mathbf{v}$$

with

$$\mathscr{K}_1 := \begin{bmatrix} \bar{K} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \mathscr{K}_2 := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \bar{K} & 0 & 0 & 0 & 0 \\ 0 & \bar{K} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \bar{K} & 0 \end{bmatrix}$$

Defining

$$\mathscr{A}_u = \mathscr{K}_1 + \mathscr{K}_2\mathscr{A} \qquad \mathscr{B}_u = \mathscr{K}_2\mathscr{B} + I \qquad \mathscr{G}_u = \mathscr{K}_2\mathscr{G}$$

enables the cost function to be expressed as

$$V_N(x_0,\mathbf{v},\mathbf{w}) = \frac{1}{2}\left|\bar{A}x_0 + \bar{B}\mathbf{v} + \bar{G}\mathbf{w}\right|_{\bar{D}}^2$$

with

$$\mathscr{Q} := \operatorname{diag}\left(\operatorname{diag}_{N-1}(Q), P_f\right) \qquad \mathscr{R} := \operatorname{diag}_N(R)$$

$$\bar{D} = \operatorname{diag}(Q, \mathscr{Q}, \mathscr{R})$$

$$\bar{A} = \begin{bmatrix} I \\ \mathscr{A} \\ \mathscr{A}_u \end{bmatrix} \qquad \bar{B} = \begin{bmatrix} 0 \\ \mathscr{B} \\ \mathscr{B}_u \end{bmatrix} \qquad \bar{G} = \begin{bmatrix} 0 \\ \mathscr{G} \\ \mathscr{G}_u \end{bmatrix}$$

The disturbance attenuation constrained optimization problem is

$$\min_{\mathbf{v}} \max_{\mathbf{w}} \frac{V_N(x_0,\mathbf{v},\mathbf{w})}{\alpha} \qquad \text{s.t. } \|\mathbf{w}\|_2 = \alpha \tag{4}$$

III. OPTIMIZATION METHODS

We present two equivalent solution methods to the disturbance attenuation problem (4): the explicit inner maximization method and the saddle point method.

A. Solution Methods Propositions

We introduce two propositions required to derive the optimal solutions to the disturbance attenuation problem. The proofs are given in the Appendix. Denote the eigenvalue decomposition of $\bar{G}'\bar{D}\bar{G}$ by

$$\bar{G}'\bar{D}\bar{G} = UTU' = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \mu_1 I_p & \\ & T_2 \end{bmatrix} \begin{bmatrix} U_1' \\ U_2' \end{bmatrix}$$

where U is orthogonal, T and T_2 are diagonal, and μ_1 is the largest eigenvalue of $\bar{G}'\bar{D}\bar{G}$ with multiplicity $p \leq n$. Note that $\left|\bar{G}'\bar{D}\bar{G}\right| = \mu_1$. First, we present Proposition 1, which is an important solution step for the explicit inner maximization solution method.

Proposition 1 (Inner max problem)**.** *Consider the optimization problem*

$$\max_{\mathbf{w}} f(\mathbf{w},\lambda) = (\bar{G}\mathbf{w} - d)'\bar{D}(\bar{G}\mathbf{w} - d) - \lambda(\mathbf{w}'\mathbf{w} - 1) \quad (5)$$

where $d = -(\bar{B}\mathbf{v} + \bar{A}x_0)$. Optimization problem (5) *has optimal value function*

$$f(\mathbf{w}^0(\lambda),\lambda) = \begin{cases} \infty, & \text{if } \lambda < \mu_1 \\ \infty, & \text{if } \lambda = \mu_1 \text{ and } U_1'\bar{G}'\bar{D}d \neq 0 \\ \mu_1 + d'\left[\bar{D} - \bar{D}\bar{G}(\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+ \bar{G}'\bar{D}\right]d, & \\ \qquad \text{if } \lambda = \mu_1 \text{ and } U_1'\bar{G}'\bar{D}d = 0 & \\ \lambda + d'\left[\bar{D} - \bar{D}\bar{G}(\bar{G}'\bar{D}\bar{G} - \lambda I)^{-1} \bar{G}'\bar{D}\right]d, & \\ \qquad \text{if } \lambda > \mu_1 & \end{cases} \quad (6)$$

and the solution for the cases of finite f^0 are

$$\mathbf{w}^0(\lambda) = \begin{cases} (\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+ \bar{G}'\bar{D}d + U_1\beta_1, & \\ \qquad \text{if } \lambda = \mu_1 \text{ and } U_1'\bar{G}'\bar{D}d = 0 & \\ (\bar{G}'\bar{D}\bar{G} - \lambda I)^{-1}\bar{G}'\bar{D}d, & \text{if } \lambda > \mu_1 \end{cases} \quad (7)$$

where $\beta_1 \in \mathbb{R}^p$ is an arbitrary vector.

Proposition 2 is necessary for the saddle point solution method.

Proposition 2. *Consider the quadratic function $f(\cdot) : \mathbb{R}^{n+m} \to \mathbb{R}$*

$$f(\mathbf{u},\mathbf{w}) := \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix}' \underbrace{\begin{bmatrix} \bar{G}'\bar{D}\bar{G} - \lambda I & -\bar{G}'\bar{D}\bar{B} \\ -\bar{B}'\bar{D}\bar{G} & \bar{B}'\bar{D}\bar{B} \end{bmatrix}}_{M} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix} + \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix}' \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

with $b \in \mathbb{R}^{n+m}$.

1) *A solution to $\max_{\mathbf{w}} \min_{\mathbf{u}} f$ exists if and only if $b \in R(M)$. Similarly, a solution to $\min_{\mathbf{u}} \max_{\mathbf{w}} f$ exists if and only if $b \in R(M)$.*
2) *For $b \in R(M)$, strong duality holds so that*

$$\min_{\mathbf{u}} \max_{\mathbf{w}} f(u,w) = \max_{\mathbf{w}} \min_{\mathbf{u}} f(\mathbf{u},\mathbf{w}) = f(\mathbf{u}^*,\mathbf{w}^*)$$

where $(\mathbf{u}^*,\mathbf{w}^*)$ are saddle points of the function f, satisfying

$$\begin{bmatrix} \mathbf{u}^* \\ \mathbf{w}^* \end{bmatrix} \in -M^+ b + N(M) \qquad f(\mathbf{u}^*,\mathbf{w}^*) = -b'M^+ b$$

and $df(\mathbf{u},\mathbf{w})/d(\mathbf{u},\mathbf{w}) = 0$ at $(\mathbf{u}^*,\mathbf{w}^*)$.

B. Explicit Inner Maximization Solution Method

We expand upon previous results on robust least squares design as described by Sayed et al. [9], and define an unconstrained optimization problem equivalent to (4) using Lagrange multipliers, $\tilde{\lambda} \in \mathbb{R}$

$$\min_{\mathbf{v}} \max_{\mathbf{w}} \min_{\tilde{\lambda}} L(\tilde{\lambda},\mathbf{v},\mathbf{w};x_0)$$

where the Lagrangian function is defined as

$$L(\tilde{\lambda},\mathbf{v},\mathbf{w};x_0) := \frac{V_N(\mathbf{v},\mathbf{w};x_0)}{\alpha} - \tilde{\lambda}(\|\mathbf{w}\|_2 - \alpha)$$

or equivalently

$$L(\lambda,\mathbf{v},\mathbf{w};x_0) := \frac{1}{\alpha^2}(\alpha V_N(\mathbf{v},\mathbf{w};x_0) - \lambda(\|\mathbf{w}\|_2 - \alpha))$$

where $\lambda = \alpha^2\tilde{\lambda}$. Strong duality holds for the operators $\max_{\mathbf{w}}$ and $\min_{\lambda}$

$$\max_{\mathbf{w}} \min_{\lambda} L(\lambda,\mathbf{v},\mathbf{w};x_0) = \min_{\lambda} \max_{\mathbf{w}} L(\lambda,\mathbf{v},\mathbf{w};x_0)$$

Where

$$\min_{\mathbf{v}} \min_{\lambda} \max_{\mathbf{w}} L(\lambda,\mathbf{v},\mathbf{w};x_0) \quad (8)$$

is the dual Lagrangian problem statement. Strong duality allows to characterize the solutions of the original optimization problem through the dual Lagrangian problem. As proved in the Appendix, optimization problem (8), which is equivalent to

$$\min_{\lambda} \min_{\mathbf{v}} \max_{\mathbf{w}} L(\lambda,\mathbf{v},\mathbf{w};x_0) \quad (9)$$

is unbounded for $\lambda < \alpha\mu_1$, and structures for the control variable sequence $\mathbf{v}$ and the disturbance sequence $\mathbf{w}$ as a function of the Lagrange multiplier λ, $\mathbf{v}^0(\lambda)$ and $\mathbf{w}^0(\lambda)$, are derived. Therefore, the disturbance attenuation optimization problem using the dual Lagrangian problem is reduced to a scalar minimization over the variable λ

$$\min_{\lambda \geq \alpha\mu_1} L(\lambda,\mathbf{v}(\lambda),\mathbf{w}(\lambda);x_0) \quad (10)$$

Where the optimal solutions are:

1) $\lambda^0 = \alpha\mu_1$

$$\mathbf{v}^0 = \mathscr{H}x_0$$
$$\mathbf{w}^0 = \mathscr{J}x_0 + U_1\mathbf{q}$$

Where $\mathbf{q}$ is an arbitrary vector. The term $U_1\mathbf{q}$ is necessary to satisfy the constraint $\|\mathbf{w}\|_2^2 = \alpha$, as shown in Proposition 1. Thus the optimal control input at time $i \in \mathbb{I}_{0:N}$ is

$$u^0(t,\lambda^0;x_0) = \bar{K}x(t) + H(t)x_0$$

2) $\lambda^0 > \alpha\mu_1$

$$\mathbf{v}^0 = \mathscr{K}x_0$$
$$\mathbf{w}^0 = \mathscr{J}x_0$$

Thus the optimal control input at time $i \in \mathbb{I}_{0:N}$ is

$$u^0(t,\lambda^0;x_0) = \bar{K}x(t) + K(t;x_0)x_0$$

Where

$$\mathscr{H} = \begin{bmatrix} H(0) \\ H(1) \\ \vdots \\ H(N-1) \end{bmatrix} \quad \mathscr{J} = \begin{bmatrix} J(0) \\ J(1) \\ \vdots \\ J(N-1) \end{bmatrix} \quad \mathscr{K} = \begin{bmatrix} K(0) \\ K(1) \\ \vdots \\ K(N-1) \end{bmatrix}$$

The control gains $H(t)$ and $K(t;x_0)$ are obtained from deriving the solution structure for $\mathbf{v}^0(\lambda)$ for the cases $\lambda^0 = \alpha\mu_1$ and $\lambda^0 > \alpha\mu_1$, respectively. The structure of the optimal gains are

1) $\lambda^0 = \alpha\mu_1$

$$\bar{F} := (\bar{D} - \bar{D}\bar{G}(\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+\bar{G}\bar{D})$$

$$\mathscr{H} = -(\bar{B}\bar{F}\bar{B})^+((\bar{B}\bar{D}\bar{G}U_1'(U_1'\bar{G}\bar{D}\bar{B}(\bar{B}\bar{F}\bar{B})^+\bar{B}'\bar{D}\bar{G}U_1')^+ (U_1'\bar{G}'\bar{D}\bar{A} + U_1'\bar{G}\bar{D}\bar{B}(\bar{B}\bar{F}\bar{B})^+\bar{B}'\bar{F}\bar{A})) + \bar{B}'\bar{F}\bar{A})$$

$$\mathscr{J} = -(\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+\bar{G}'\bar{D}(\bar{B}\mathscr{H} + \bar{A})$$

2) $\lambda^0 > \alpha\mu_1$

$$\bar{F} := \alpha(\bar{D} - \bar{D}\bar{G}(\alpha\bar{G}'\bar{D}\bar{G} - \lambda^0 I)^{-1}\bar{G}\bar{D})$$

$$\mathscr{K} = -(\bar{B}\bar{F}\bar{B})^{-1}\bar{B}\bar{F}\bar{A}$$

$$\mathscr{J} = -\alpha(\alpha\bar{G}'\bar{D}\bar{G} - \lambda^0 I)^{-1}\bar{G}'\bar{D}(\bar{B}\mathscr{K} + \bar{A})$$

C. Saddle Point Solution Method

Optimization problem (9) is strongly dual for the operators $\min_{\mathbf{v}}$ and $\max_{\mathbf{w}}$ and the optimal solutions as a function of λ are obtained from the necessary conditions

$$\nabla_{\mathbf{v},\mathbf{w}} L(\lambda, \mathbf{v}, \mathbf{w}; x_0) = 0$$

By factoring (9)

$$L(\lambda, \mathbf{v}, \mathbf{w}; x_0) = \frac{1}{\alpha^2}\left(\begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix}' M \begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix} - 2\begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix}' b + \alpha x_0'\bar{E}x_0 + \alpha\lambda \right)$$

$$M := \begin{bmatrix} \alpha\bar{G}'\bar{D}\bar{G} - \lambda I & -\alpha\bar{G}'\bar{D}\bar{B} \\ -\alpha\bar{B}'\bar{D}\bar{G} & \alpha\bar{B}'\bar{D}\bar{B} \end{bmatrix}$$

$$b := \begin{bmatrix} \bar{G}' \\ -\bar{B}' \end{bmatrix} \alpha\bar{D}\bar{A}x_0 \qquad \bar{E} := \bar{A}'\bar{D}\bar{A}$$

And solving for the necessary conditions, we obtain

$$M \begin{bmatrix} \mathbf{w}(\lambda) \\ \mathbf{v}(\lambda) \end{bmatrix}^0 = b$$

And from Proposition 2 in the Appendix the solution is

$$\begin{bmatrix} \mathbf{w}(\lambda) \\ \mathbf{v}(\lambda) \end{bmatrix}^0 = M^+ b + N(M) \tag{11}$$

Statement (10) is solved using the functions $\mathbf{w}(\lambda)$ and $\mathbf{v}(\lambda)$ obtained in (11), and the optimal solutions for optimization problem (9), $\mathbf{v}^0$ and $\mathbf{w}^0$, which are equivalent to the ones obtained through the explicit inner maximization solution method, are recovered. The saddle point solution method is valid for $\lambda^0 \geq \alpha\mu_1$. When $\lambda^0 > \alpha\mu_1$, the pseudoinverse, M^+, becomes an inverse, M^{-1}, as the matrix M is always nonsingular in this case. When $\lambda^0 = \alpha\mu_1$, the nullspace $N(\alpha\bar{G}'\bar{D}\bar{G} - \alpha\mu_1 I)$ is necessary to satisfy the constraint $\|\mathbf{w}\|_2^2 = \alpha$

$$\mathbf{w}^0 \in \mathbf{w}^0(\lambda^0) + N(\alpha\bar{G}'\bar{D}\bar{G} - \alpha\mu_1 I)$$

as shown in the proof for Proposition 2 in the Appendix.

Regardless of the chosen solution method, the optimal controller is linear and independent of the parameter α, when $\lambda^0 = \alpha\mu_1$. The optimal controller is instead nonlinear and depends on the parameter α, when $\lambda^0 > \alpha\mu_1$. Furthermore, if $\lambda^0 = \alpha\mu_1$, the controller is equivalent to the H_∞ optimal control.

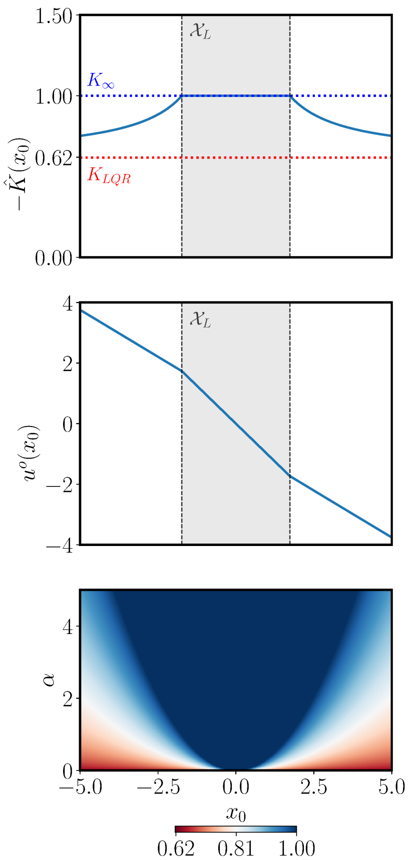

Fig. 1. Top. Optimal control gain $\hat{K}$ as a function of x_0, for $\alpha = 1$. Middle. Optimal control input $u^0(0, \lambda^0)$ as a function of x_0, for $\alpha = 1$. Bottom. Contour plot of $\hat{K}$ as a function of x_0 and α.

IV. SOLUTION TO THE DISTURBANCE ATTENUATION PROBLEM

Optimization problem (9) features two solution regions in the space of the initial states. We define the following region, $\mathscr{X}_L$ where the optimal solution is $\lambda^0 = \alpha\mu_1$

$$\mathscr{X}_L := \{x_0 \in \mathbb{R}^N : x_0'\mathscr{J}'\mathscr{J}x_0 < \alpha\} \tag{12}$$

Where the gain $\mathscr{J}$ is computed for $\lambda = \alpha\mu_1$, and is obtained using either solution method presented in Section III. When the arbitrary gain $\bar{K}$ is chosen to be K_∞, the optimal control gain from the solution of the H_∞ control problem, then (12) characterizes a region with an optimal linear control, which

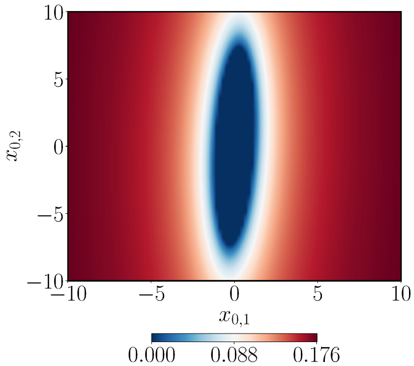

Fig. 2. Contour plot of the relative difference between the optimal control gain $|\hat{K}|$ and the optimal H_∞ control gain $|K_\infty|$, for $\alpha = 1$.

is equivalent to the optimal H_∞ control. Indeed by choosing $\bar{K} = K_\infty$, we obtain

$$\mathbf{v}^0 = \mathscr{H} x_0, \qquad \mathscr{H} = 0$$

And thus the optimal control input, which is linear and independent of x_0, at time $i \in \mathbb{I}_{0:N}$ is

$$u^0(t;\lambda^0) = \bar{K}x(t) = K_\infty x(t), \quad x_0 \in \mathscr{X}_L$$

The region $\mathscr{X}_{NL} := \mathbb{R}^n \setminus \mathscr{X}_L$, where the optimal solution is $\lambda^0 > \alpha\mu_1$, features instead a nonlinear optimal controller, which converges to the optimal LQR control when the initial state x_0 becomes large, and is defined as

$$\mathscr{X}_{NL} := \{x_0 \in \mathbb{R}^N : x_0' \mathscr{J}' \mathscr{J} x_0 \geq \alpha, \} \tag{13}$$

Where the gain $\mathscr{J}$ is computed for $\lambda = \alpha\mu_1$. Both regions $\mathscr{X}_L$ in (12) and $\mathscr{X}_{NL}$ in (13) are characterizable a priori. Once the region is known, then solution methods from Section III are used to retrieve the optimal control solution, $\mathbf{v}^0$, to implement in potential applications. While the optimization methods are formulated within a finite horizon framework, choosing a large horizon length provides open-loop convergence to the optimal solutions at the first time step, i.e., $t = 0$. A numerical scalar example, based on [10, p. 92], for the following system

$$A = 1 \quad B = 1 \quad R = 1 \quad Q = 1$$
$$-K_\infty = 1 \quad -K_{LQR} = 0.62$$

demonstrates the overall nonlinear behavior of the optimal control gain, $\hat{K} = K_\infty + K(0;x_0)$ for $x_0 \in \mathscr{X}_{NL}$ or $\hat{K} = K_\infty + H(0)$ for $x_0 \in \mathscr{X}_L$, and control input, $u^0(0;\lambda^0) = \hat{K}x_0$, and the transition from region $\mathscr{X}_L$ to region $\mathscr{X}_{NL}$, as illustrated in the top and middle plots of Figure 1. The dependence on the parameter α of the optimal control gain, $\hat{K}$, is shown in the bottom plot of Figure 1. Within region $\mathscr{X}_L$, the optimal linear control is independent of the parameter α. As expected, for $\alpha = 0$, the optimal LQR control is recovered. A multidimensional numerical example for the following system

$$A = \begin{bmatrix} 2 & 0.5 \\ 1.5 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0.5 \\ 1 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is used to compute the relative difference between $|\hat{K}|$ and $|K_\infty|$ to clearly illustrate the elliptical nature of the boundary of region $\mathscr{X}_L$, $\partial\mathscr{X}_L$, where the optimal control is linear and equivalent to $|K_\infty|$, as shown in Figure 2.

V. CONCLUSIONS

This paper expands upon the previous work on H_∞ optimal control and demonstrates the nonlinearity of the optimal feedback control of the disturbance attenuation problem for linear dynamical systems, providing a unified framework that spans from robust H_∞ control to LQR control as a function of the initial state compared to the magnitude of the disturbance. Future work includes 1) applying the proposed solution methods to a closed-loop framework, i.e., robust model predictive control, and investigate relevant applications and 2) develop a state estimator for the proposed solution methods to account for imperfect measurements.

APPENDIX

Proof of Proposition 1.

Proof. Perform an invertible coordinate transformation $\mathbf{w} = U\beta, \beta = U'\mathbf{w}$, and we maximize over β in place of $\mathbf{w}$. The objective function is

$$f(\mathbf{w},\lambda) = \mathbf{w}'(\bar{G}'\bar{D}\bar{G} - \lambda I)\mathbf{w} - 2d'\bar{D}\bar{G}\mathbf{w} + d'\bar{D}d + \lambda$$
$$f(\alpha,\lambda) = (\mu_1 - \lambda)\beta_1'\alpha_1 + \beta_2'(M_2 - \lambda I)\beta_2 - 2(\mathbf{y}_1'\beta_1 + \mathbf{y}_2'\beta_2) + d'\bar{D}d + \lambda$$

with vector $\mathbf{y} = U'\bar{G}'\bar{D}d$. We have the following four cases

1) $\lambda < \mu_1$. Choose $\beta_1 \to \infty$ and $\beta_2 = 0$ to obtain $f \to \infty$, which is the first result.
2) $\lambda = \mu_1, U_1'\bar{G}'\bar{D}d = \mathbf{y}_1 \neq 0$. We then have
$$f(\beta,\lambda) = \beta_2'(T_2 - \lambda I)\beta_2 - 2(\mathbf{y}_1'\beta_1 + y_2'\beta_2) + d'Dd + \lambda$$
Choose $\beta_1 = -\rho y_1$, $\beta_2 = 0$ and let $\rho \to \infty$ to obtain $f \to \infty$, which is the second result.
3) $\lambda = \mu_1$, $U_1'\bar{G}'\bar{D}d = \mathbf{y}_1 = 0$. We then have
$$f(\beta,\lambda) = \beta_2'(T_2 - \mu_1 I)\beta_2 - 2\mathbf{y}_2'\beta_2 + d'Dd + \lambda$$
Matrix $(T_2 - \mu_1 I)$ is invertible, so the optimal value is $\beta_2^0 = (T_2 - \mu_1 I)^{-1}\mathbf{y}_2$, and β_1 is arbitrary. Converting back to $\mathbf{w}$ gives
$$\mathbf{w}^0 = U_1\beta_1 + U_2(T_2 - \mu_1 I)^{-1}U_2'G'Dd$$
$$\mathbf{w}^0 = U_1\beta_1 + (\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+\bar{G}'\bar{D}d$$
where we have used the pseudo-inverse in place of the SVD expression. Substituting this result into the expression for $f(\mathbf{w},\lambda)$ above gives
$$f^0 = \mu_1 + d'\left[\bar{D} - \bar{D}\bar{G}(\bar{G}'\bar{D}\bar{G} - \mu_1 I)^+\bar{G}'\bar{D}\right]d$$
and we have established the third case.
4) $\lambda > \mu_1$. The matrix $T - \lambda I$ is now invertible and the optimal solution is
$$\beta^0 = (T - \lambda I)^{-1}\mathbf{y} \quad \mathbf{w}^0 = (\bar{G}'\bar{D}\bar{G} - \lambda I)^{-1}\bar{G}'\bar{D}d$$

and substituting this result into the expression for $f(\mathbf{w},\lambda)$ gives

$$f^0 = \lambda + d'\left[\bar{D} - \bar{D}\bar{G}(\bar{G}'\bar{D}\bar{G} - \lambda I)^{-1}\bar{G}'\bar{D}\right]d$$

$\square$

Proof of Proposition 2.

Proof. First we show that $b \in R(M)$ implies existence of solutions to both minmax and maxmin problems. It is convenient to organize the terms in f two ways

$$f(\mathbf{u},\mathbf{w}) = \frac{1}{2}\mathbf{u}'M_{22}\mathbf{u} + \mathbf{u}'(b_2 + M_{12}'\mathbf{w}) - \frac{1}{2}\mathbf{w}'M_{11}\mathbf{w} + \mathbf{w}'b_1 \tag{14}$$

$$= -\frac{1}{2}\mathbf{w}'M_{11}\mathbf{w} + \mathbf{w}'(b_1 + M_{12}\mathbf{u}) + \frac{1}{2}\mathbf{u}'M_{22}\mathbf{u} + \mathbf{u}'b_2 \tag{15}$$

with $M_{11} = \bar{G}'\bar{D}\bar{G} - \lambda I$, $M_{12} = -\bar{G}'\bar{D}\bar{B}$, and $M_{22} = \bar{B}'\bar{D}\bar{B}$

a) Maxmin exists: Consider the inner problem $\min_{\mathbf{u}} f(\mathbf{u},\mathbf{w})$ using (14). The optimal solution satisfies $df/d\mathbf{u} = 0$, giving

$$M_{22}\underline{\mathbf{u}}^0 + M_{12}'\mathbf{w} + b_2 = 0 \tag{16}$$

which, when feasible, implicitly defines a function (point-to-set map) $\underline{\mathbf{u}}^0(\mathbf{w})$ that is the optimal solution of the inner problem as a function of $\mathbf{w}$. Next we solve the outer maximization $\max_z V(z,\underline{y}^0(z))$. To take the derivative we use the chain rule

$$\frac{d}{d\mathbf{w}}d(\mathbf{w},\underline{\mathbf{u}}^0(\mathbf{w})) = \left(\partial f/\partial \mathbf{w}\right)_{\underline{\mathbf{u}}^0} + j'\left(\partial f/\partial \underline{\mathbf{u}}^0\right)_{\mathbf{w}}$$

where $j = (\partial \underline{\mathbf{u}}^0(\mathbf{w})/\partial \mathbf{w})$ is the Jacobian of the vector-valued function $\underline{\mathbf{u}}^0(\mathbf{w})$. Taking the partial derivatives using expression (15) for the first and (14) for the second gives

$$\left(\partial f/\partial \mathbf{w}\right)_{\underline{\mathbf{u}}^0} = -M_{11}\mathbf{w} + M_{12}\underline{\mathbf{u}}^0 + b_1$$

$$\left(\partial f/\partial \underline{\mathbf{u}}^0\right)_{\mathbf{w}} = M_{22}\underline{\mathbf{u}}^0 + M_{12}'\mathbf{w} + b_2$$

Setting $df/d\mathbf{w} = 0$ then gives a linear equation for the optimal $\mathbf{w}$, $-M_{11}\mathbf{w}^0 + M_{12}\underline{\mathbf{u}}^0 + b_1 = 0$, and combining with (16) gives the set of equations for the maxmin solution

$$\begin{bmatrix} -M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix}\begin{bmatrix} \mathbf{w}^0 \\ \underline{\mathbf{u}}^0 \end{bmatrix} = -\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad M\begin{bmatrix} \mathbf{w}^0 \\ \underline{\mathbf{u}}^0 \end{bmatrix} = -b$$

Notice that this set of equations is feasible since $b \in R(M)$, and we have the solution

$$\begin{bmatrix} \mathbf{w}^0 \\ \underline{\mathbf{u}}^0 \end{bmatrix} \in -M^+b + N(M)$$

which shows that $\mathbf{w}^0 = \mathbf{w}^*$. If we want to evaluate the corresponding $\underline{\mathbf{u}}^0$ values, then we take the optimal $\mathbf{w}$ and solve (16) yielding

$$\underline{\mathbf{u}}^0(\mathbf{w}^*) \in -M_{22}^+(M_{12}'\mathbf{w}^* + b_2) + N(M_{22})$$

Note that $\mathbf{u}^* \in \underline{\mathbf{u}}^0(\mathbf{w}^*)$.

b) Minmax exists.: Proceeding analogously, we consider the inner problem $\max_{\mathbf{w}} f(\mathbf{u},\mathbf{w})$ using expression (15), and set $df/d\mathbf{w} = 0$ to obtain

$$-M_{11}\overline{\mathbf{w}}^0 + M_{12}\mathbf{u} + b_1 = 0 \tag{17}$$

where we now have an implicit function of the optimal $\mathbf{w}$, $\overline{\mathbf{w}}^0(\mathbf{u})$ as a function of parameter $\mathbf{u}$. Then we solve outer $\min_{\mathbf{u}} f(\overline{\mathbf{w}}^0(\mathbf{u}),\mathbf{u})$ by setting $df/d\mathbf{u} = 0$. Proceeding as above, we obtain a linear equation for the optimal $\mathbf{u}$, $M_{22}\mathbf{u}^0 + M_{12}'\overline{\mathbf{w}}^0 + b_2 = 0$. Combining with (17) we obtain an identical set of equations for the minmax problem

$$\begin{bmatrix} -M_{11} & M_{12} \\ M_{12}' & M_{22} \end{bmatrix}\begin{bmatrix} \overline{\mathbf{w}}^0 \\ \mathbf{u}^0 \end{bmatrix} = -\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad M\begin{bmatrix} \overline{\mathbf{w}}^0 \\ \mathbf{u}u^0 \end{bmatrix} = -b$$

The equations are feasible since $b \in R(M)$, and we have obtained the solution for the minmax problem

$$\begin{bmatrix} \overline{\mathbf{w}}^0 \\ \mathbf{u}^0 \end{bmatrix} \in -M^+b + N(M)$$

which shows that $\mathbf{u}^0 = \mathbf{u}^*$. As before we can find the corresponding $\overline{\mathbf{w}}^0$ values from (17)

$$\overline{\mathbf{w}}^0(\mathbf{u}^*) \in M_{22}^+(M_{12}\mathbf{u}^* + b_1) + N(M_{11})$$

and we note that $\mathbf{w}^* \in \overline{\mathbf{w}}^0(\mathbf{u}^*)$. So for $b \in R(M)$, we have established strong duality. The optimal value is then

$$f(\mathbf{u}^*,\mathbf{w}^*) = \frac{1}{2}b'M^+MM^+b - b'M^+b = -\frac{1}{2}b'M^+b$$

Finally, if $b \notin R(M)$, we have no solution to (16) in the maxmin problem for which $df(\mathbf{w},\underline{\mathbf{u}}^0(\mathbf{w}))/d\mathbf{w} = 0$, so the maxmin does not have a solution. Analogously, there is no solution to (17) for which $df(\overline{\mathbf{w}}^0(\mathbf{u}),\mathbf{u})/d\mathbf{u} = 0$, so the minmax problem also does not have a solution, and Proposition 2 has been established. $\square$

REFERENCES

[1] I. Postlethwaite, M. C. Turner, and G. Herrmann, "Robust control applications," *Annual Reviews in Control*, vol. 31, no. 1, pp. 27–39, 2007.

[2] P. Coustal and J.-M. Michelin, "Industrial application of an H-infinity design method for flexible structures," *IEEE Control Systems Magazine*, vol. 14, no. 4, pp. 49–54, 1994.

[3] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Auto. Cont.*, vol. 26, no. 2, pp. 301–320, Apr 1981.

[4] K. Glover and J. C. Doyle, "State-space formulae for all stabilizing controllers that satisfy an H_∞-norm bound and relations to risk sensitivity," *Sys. Cont. Let.*, vol. 11, no. 3, pp. 167–172, 1988.

[5] T. Basar, "A dynamic games approach to controller design: Disturbance rejection in discrete time," in *Proceedings of the 28th IEEE Conference on Decision and Control,*, 1989, pp. 407–414.

[6] G. Didinsky and T. Basar, "Design of minimax controllers for linear systems with non-zero initial states under specified information structures," *Int. J. Robust and Nonlinear Control*, vol. 2, no. 1, pp. 1–30, 1992.

[7] T. Basar, "Disturbance attenuation in LTI plants with finite horizon: Optimality of nonlinear controllers," *Sys. Cont. Let.*, vol. 13, no. 3, pp. 183–191, 1989.

[8] E. C. Titchmarsh, *Introduction to the theory of Fourier integrals*, 3rd ed. New York, NY: Chelsea Publishing Company, 1986.

[9] A. H. Sayed, V. H. Nascimento, and F. A. Cipparrone, "A regularized robust design criterion for uncertain data," *SIAM J. Matrix Anal. and Appl.*, vol. 23, no. 4, pp. 1120–1142, 2002.

[10] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Boston: Birkhäuser, 1995.