GENERIC COL- ORIZED JOUR- NAL, VOL. XX, NO. XX, XXXX 2017 1

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

Robust Data-Driven Control of Discrete-Time Linear Systems with Errors in Variables

Jared Miller, *Member, IEEE*, Tianyu Dai, *Member, IEEE*, and Mario Sznaier *Fellow, IEEE*

***Abstract*— This work presents a Sum of Squares (SOS) based framework to perform data-driven stabilization and robust control tasks on discrete-time linear systems where the full-state observations are corrupted by ℓ^∞ bounded input, measurement, and process noise (error in variable setting). Certificates of full-state-feedback robust performance, superstabilization or quadratic stabilization of all plants in a consistency set are provided by solving a feasibility program formed by polynomial nonnegativity constraints. Under mild compactness and data-collection assumptions, SOS tightenings in rising degree will converge to recover the true worst-case optimal ℓ^∞ (extended) superstabilizing controllers. With some conservatism, quadratically stabilizing controllers with certified $\mathcal{H}_2$ performance bounds can also be found. The performance of this SOS method is improved through the application of a theorem of alternatives while retaining tightness, in which the unknown noise variables are eliminated from the consistency set description. This SOS feasibility method is extended to provide worst-case-optimal robust controllers under $\mathcal{H}_2$ control costs.**

***Index Terms*— Data-Driven Control, Robust Control, Optimization, Linear Matrix Inequality, Sum of Squares**

I. Introduction

The traditional "data to control" pipeline involves a two step process: (i) identifying a nominal model of the plant and an associated bounded uncertainty description from noisy, finite data; and (ii) designing a robust controller that guarantees worst case performance over all admissible uncertainty. Both of these steps can lead to generically NP-hard problems, necessitating the use of computationally expensive, potentially conservative relaxations. As an alternative, the last few years have seen renewed interest in direct Data-Driven Control (DDC), a group of methods that sidestep system-identification in order to design controllers that guarantee worst case performance for all possible plants which are consistent with observed data. In this paper, we seek to extend DDC to the case where the measurements of both the state and control action are corrupted by noise (the so-called Error in Variables (EIV) setting). Specifically, we consider discrete-time linear systems with states $x_t \in \mathbb{R}^n$ and inputs $u_t \in \mathbb{R}^m$ for which measured data, corrupted by noise up to a finite time horizon of T is available as $\mathcal{D} = \{\hat{u}_t, \hat{x}_t\}_{t=1}^T$. The system includes ℓ^∞-bounded full-state measurement noise Δx_t, input noise Δu_t, and process noise w_t to form the model

$$x_{t+1} = Ax_t + Bu_t + w_t \tag{1a}$$
$$\hat{x}_t = x_t + \Delta x_t, \quad \hat{u}_t = u_t + \Delta u_t. \tag{1b}$$

It is desired to find a constant matrix $K \in \mathbb{R}^{m\times n}$ generating a state-feedback law $u = Kx$ such that the closed loop system achieves robust performance for all plants (A, B) that could have generated the data in $\mathcal{D}$. In particular we will consider the case where performance is given in terms of a suitable norm of the closed loop system. A large portion of existing DDC work involves noiseless systems or sets $\Delta x, \Delta u = 0$, considering only process noise w (allowing for efficient convex optimization). [1]–[11]. The work in [3] defines a notion of 'data informativity,' demonstrating that the assumptions required for a data-driven stabilization task are less restrictive than the assumptions needed to perform system identification. Semidefinite Program (SDP)-based methods for nonconservative DDC under ℓ^2-bounded process noise (with $\Delta x, \Delta u = 0$) include using a matrix S-Lemma [4], Petersen's Lemma [5], multipliers learned from data [6], and Lyapunov-Metzler (bilinear) inequalities for switched systems [7].

ℓ^∞ bounded process noise was considered in [1], [2], [6], [8], [9], [11]. This scenario arises for instance from error propagation of finite-difference approximations when computing derivatives and sampling. An advantage of ℓ^∞ process noise as compared to ℓ^2 process noise is that multiple datasets $\mathcal{D}$ with differing time horizons can be concatenated without scaling or shifting the noise effects. The work in [6] briefly mentions adaptation for the ℓ^∞-bounded process noise case, while the computational complexity of ℓ^∞-bounded stabilization increases in an exponential manner with the number of measurements in $\mathcal{D}$ (T). A lower complexity, Sum of

This paper was submitted for review on August 3, 2023, was revised on February 14, 2024, was conditionally accepted on March 31, 2024, and was accepted on August 8, 2024. J. Miller, T. Dai, and M. Sznaier were partially supported by NSF grants CNS–2038493 and CMMI–2208182, AFOSR grant FA9550-19-1-0005, ONR grant N00014-21-1-2431, and the Sentry DHS Center of Excellence under Award 22STESE00001-03. J. Miller was in part supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States, AFOSR grant FA9550-19-1-0005, the AFOSR International Student Exchange Program , and the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

J. Miller is with the Automatic Control Laboratory (IfA), Department of Information Technology and Electrical Engineering (D-ITET), ETH Zürich, Physikstrasse 3, 8092, Zürich, Switzerland (jarmiller@control.ee.ethz.ch)

T. Dai is with The MathWorks, Inc., 1 Apple Hill Drive, Natick, MA 01760 USA (tdai@mathworks.com)

J. Miller and M. Sznaier are with the Robust Systems Lab, ECE Department, Northeastern University, Boston, MA 02115 USA (msznaier@coe.neu.edu).

Squares (SOS)-based, approach addressing the same problem can be found in [9]. Alternatively, complexity can be reduced through the use of the (potentially conservative) notion of superstability [12]. Work along these lines can be found in [1], [2], [8]. Finally, the work in [11] uses Linear Programs (LPs) to perform positive stabilization of linear systems under ℓ^∞-bounded process noise.

Compared to the scenarios above, the EIV-aware control setting addressed in this paper is substantially more involved and has been considered in relatively few papers. The set of plants (A, B) consistent with ℓ^∞-bounded process noise forms a polytope and robust performance can be achieved, in principle, by solving a collection of vertex LMIs [8]. On the other hand, expressing the model (1) purely in terms of observations $\mathcal{D}$ and noise processes $(\Delta x, \Delta u, w)$ leads to:

$$\hat{x}_{t+1} - \Delta x_{t+1} = A(\hat{x}_t - \Delta x_t) + B(\hat{u}_t - \Delta u_t) + w_t. \quad (2)$$

Equation (2) involves multiplications $(A\Delta x_t, B\Delta u_t)$ between unknown variables inside the description of the set (A, B) of data-consistent plants. Thus, the plants consistent with EIV are generically contained in a non-convex region [13]. Further, as illustrated in Section VII, using process noise to characterize this region can lead to controllers that fail to stabilize the actual plant.

Prior work on set membership systems identification in an EIV setting includes [14]–[16], which sought to overbound the consistency set. In the context of DDC, measurement noise has been considered in [17], [18]. [17] considers ℓ^2 bounded measurement noise and provides a Linear Matrix Inequality (LMI) based synthesis procedure guaranteed to stabilize the plant provided that the signal-to-noise ratio is sufficiently large and the noise sufficiently small. However, as shown in that paper, these conditions can sometimes be overly conservative. The work in [18] handles white measurement noise through the introduction of an auxiliary slack variable, to ensure feasibility of the design constraints, and regularization. EIV scenarios with ℓ^2 bounded noise have been considered in [19], [20], both posted after the present paper was submitted, providing necessary and sufficient conditions for data-driven stabilization when the noise is characterized by a single quadratic constraint. In the case of multiple constraints, such as when it is desired to impose instantaneous bounds on the noise, these conditions are only sufficient. To the best of our knowledge, the only paper addressing DDC in an ℓ^∞ EIV setting is [21]. The present paper continues this line of work. It formulates control of all consistent plants as a Polynomial Optimization Problem (POP), which is approximated by a converging sequence of SDPs through SOS methods [22].

A theorem of alternatives based on robust SDPs is used to reduce the complexity of the generated SOS programs by eliminating the noise variables $(\Delta x, \Delta u, w)$. Its contributions are:

- Formulation of data driven robust performance under ℓ^∞ EIV as a polynomial optimization problem
- Application of SOS methods to recover constant state-feedback controllers guaranteeing worst case $(\ell^\infty, \mathcal{H}_2)$ performance of all plants consistent with the recorded data
- Simplification of SOS programs by using a Theorem of Alternatives to eliminate affine-dependent noise variables
- Analysis of computational complexity of SOS programs

Sections of this work were presented at the 61st Conference on Decision and Control [21]. New content in this journal version as compared to the conference paper includes

- Robust ℓ^∞ and $\mathcal{H}_2$ performance and their connection with extended superstabilization and quadratic stabilization
- Noting that a certificate function for superstability can be Δx-independent
- Derivation and application of the matrix Theorem of Alternatives ensuring Positive Definiteness
- Proofs of continuity, polynomial approximability, and convergence
- Further detail about the combination of measurement, input, and process noise

This paper is laid out as follows: Section II introduces preliminaries such as acronym definitions, notation, stability conditions for classes of linear systems, and SOS methods. Section III creates a Basic Semialgebraic (BSA) description of the consistency set of plants compatible with measurement-noise-corrupted data, and formulates SOS algorithms to recover worst case ℓ^∞ or $\mathcal{H}_2$ controllers. Section IV reduces the computational complexity of these SOS programs by eliminating the affine-dependent measurement noise variables through a Theorem of Alternatives. Section V quantifies this reduction in computational complexity by analyzing the size and multiplicities of Positive Semidefinite (PSD) matrices involved in these SOS methods. Section VI extends the SOS formulations to problems with measurement, input, and process noise. Section VII demonstrates the SOS based algorithms on a set of examples. Section VIII concludes the paper. Appendix I proves that multiplier functions for the Alternatives program may be chosen to be continuous. Appendix II builds on this continuity result and proves that the multiplier functions may also be chosen to be symmetric-matrix-valued polynomials.

II. Preliminaries

A. Acronyms/Initialisms

BSA	Basic Semialgebraic
DDC	Data-Driven Control
EIV	Error in Variables
LMI	Linear Matrix Inequality
LP	Linear Program
PMI	Polynomial Matrix Inequality
POP	Polynomial Optimization Problem
PSD	Positive Semidefinite
PD	Positive Definite
SDP	Semidefinite Program
SOS	Sum of Squares
WSOS	Weighted Sum of Squares

B. Notation

The set of real numbers is $\mathbb{R}$, its n-dimensional vector space is $\mathbb{R}^n$, and its n-dimensional nonnegative real orthant is $\mathbb{R}^n_+$. The set of natural numbers is $\mathbb{N}$, and the subset of natural numbers between 1 and N is $1..N$.

The set of $m \times n$ matrices with real entries is $\mathbb{R}^{m\times n}$. The transpose of a matrix Q is Q^T, and the subset of $n \times n$ symmetric matrices satisfying $Q^T = Q$ is $\mathbb{S}^n$. The square identity matrix is $I_n \in \mathbb{S}^n$. The rectangular identity matrix $I_{n\times m}$ is a matrix whose main diagonal has values of 1 with all other entries equal to zero (consistent with MATLAB's eye(n, m) function). The inverse of a matrix $Q \in \mathbb{R}^{n\times n}$ is Q^{-1}, and the inverse of its matrix transpose is Q^{-T}. The trace of a matrix Q is $\text{Tr}(Q)$. The Kronecker product of two matrices A and B is $A \otimes B$. The set of real symmetric PSD matrices $\mathbb{S}^n_+$ have all nonnegative eigenvalues ($Q \succeq 0$), and its subset of Positive Definite (PD) matrices $\mathbb{S}^n_{++}$ have all positive eigenvalues ($Q \succ 0$).

ℓ^∞ denotes the space of bounded real sequences equipped with the peak norm $\|x_t\|_\infty \doteq \sup_t \max_i |x_{it}|$. Given a linear time invariant operator $\mathcal{G} : \ell^\infty \to \ell^\infty$, we denote its ℓ^∞ to ℓ^∞induced norm by $\|\mathcal{G}\|_{\ell_\infty\to\ell_\infty} := \sup_{\|w\|_\infty\leq 1} \|\mathcal{G}w\|_\infty$. In the case of operators represented by a finite matrix $M \in \mathbb{R}^{m\times n}$ its ℓ^∞ induced norm will be denoted simply by $\|M\|_\infty = \max_i \sum_{j=1}^n |M_{ij}|$. The asterisk operator $*$ may be used to fill in transposed entries of a symmetric matrix. The minimum and maximum eigenvalues of a matrix $Q \in \mathbb{S}^n$ are $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$ respectively. The elementwise division between vectors $a, b \in \mathbb{R}^n$ is $a./b$.

The set of polynomials in variable x with real coefficients is $\mathbb{R}[x]$. The degree of a polynomial $p(x) \in \mathbb{R}[x]$ is $\deg p$. The set of polynomials with degree at most d for $d \in \mathbb{N}$ is $\mathbb{R}[x]_{\leq d}$. The set of vector-valued polynomials is $(\mathbb{R}[x])^n$, and the set of matrix-valued polynomials is $(\mathbb{R}[x])^{m\times n}$. The subset of $n\times n$ symmetric-matrix-valued polynomials is $\mathbb{S}^n[x]$, and its subcone of PSD (PD) polynomial matrices is $\mathbb{S}^n_+[x]$ ($\mathbb{S}^n_{++}[x]$). The set of SOS polynomials is $\Sigma[x]$, and the set of SOS matrices of size $n \times n$ is $\Sigma^n[x] \subset \mathbb{S}^n[x]$.

The projection operator $\pi^x : (x, y) \mapsto x$ applied to a set $X \times Y$ is $\pi^x(X \times Y) = \{x \mid (x, y) \in X \times Y\}$.

C. Optimizing Performance of Discrete-Time Systems

Consider a discrete time LTI system of the form

$$\mathcal{G} : \begin{cases} x_{t+1} &= Ax_t + Bu_t + Ew_t \\ z_t &= Cx_t + Du_t \end{cases} \tag{3}$$

where w_t is a disturbance known to belong to some set $\mathcal{W}$ and the output z measures performance. The data-driven robust performance problem addressed in this paper is to find a static control law $u_t = Kx_t$ that minimizes the worst case value of a suitable function of z over all $w_t \in \mathcal{W}$ and all pairs (A, B) consistent with the observed data. A prerequisite to posing robust performance problems is formulation of robust stabilization.

1) Superstability and the ℓ^∞ induced case: The goal here is to synthesize a control law that minimizes the induced ℓ^∞ norm of the closed loop system $\mathcal{G} : w_t \to z_t$ in (3). While this problem can be solved using Linear Programming [23], these optimal solutions can have arbitrarily high order. An elegant approach to find static controllers is to use the concept of (extended) superstability to minimize an upper bound of the ℓ^∞ induced norm. For exposition in this section, we assume that $C = I$ and $D = \mathbf{0}$.

A system $x_{t+1} = Ax_t + Ew_t$ is Extended-Superstable if there exists a vector $v > 0$, a matrix X:=diag(v), and a scalar $\lambda \geq 0$ such that [24]

$$\|X^{-1}AX\|_{\ell^\infty\to\ell^\infty} \leq \lambda < 1 \tag{4}$$

In this case, an upper bound of the ℓ^∞ induced norm of the system (3) is given by [24, Theorem 2]:

$$\sup_{\|w\|_\infty\leq 1} \|z\|_\infty \leq \frac{\max(v)}{\min(v)} \frac{\|E\|_1}{(1-\lambda)}. \tag{5}$$

Finding a controller K that minimizes this ℓ^∞ can be computed through solving the following LP parameterized by $\lambda \in [0, 1)$ [24, Theorem 3]:

$$\begin{aligned}
J(\lambda) = &\min_{M,\ S,\ v,\ \beta} \quad \frac{\beta}{1-\lambda} && && \text{(6a)}\\
&1 \leq v_i \leq \beta && \forall i \in 1..n && \text{(6b)}\\
&\textstyle\sum_{j=1}^n M_{ij} \leq \lambda v_i && \forall i \in 1..n && \text{(6c)}\\
&|A_{ij}v_j + \textstyle\sum_{\ell=1}^m B_{i\ell}S_{\ell j}| \leq M_{ij} && \forall i, j \in 1..n && \text{(6d)}\\
&v \in \mathbb{R}^n_{>0},\ \beta \in \mathbb{R},\ S \in \mathbb{R}^{m\times n}, && && \text{(6e)}\\
&M \in \mathbb{R}^{n\times n}. && && \text{(6f)}
\end{aligned}$$

The ℓ^∞ gain of the system in (3) is upper-bounded by

$$\sup_{\|w\|_\infty\leq 1} \|z\|_\infty \leq \min_{\lambda\in[0,1)} J(\lambda). \tag{7}$$

When (6) is feasible at a given λ, the corresponding controller is $K = SX^{-1}$. The resulting closed-loop satisfies $\|X^{-1}(A + BK)X\|_\infty \leq \lambda < 1$. Hence it is extended-superstable. Extended superstabilization can therefore be achieved by finding a λ such that $J(\lambda) < \infty$.

The line search over the parameter $\lambda \in [0, 1)$ can be avoided by considering a more conservative upper bound obtained by setting $v = \mathbf{1}$ ($X = I$) and $\beta = 1$. This restricted program with $K = SI = S$ is:

$$\begin{aligned}
\lambda^* = &\min_{M,\ K,\ \lambda} \lambda && && \text{(8a)}\\
&\textstyle\sum_{j=1}^n M_{ij} \leq \lambda && \forall i \in 1..n && \text{(8b)}\\
&|A_{ij} + \textstyle\sum_{\ell=1}^m B_{i\ell}K_{\ell j}| \leq M_{ij} && \forall i, j \in 1..n && \text{(8c)}\\
&M \in \mathbb{R}^{n\times n},\ K \in \mathbb{R}^{m\times n} && && \text{(8d)}\\
&\lambda \in [0, 1). && && \text{(8e)}
\end{aligned}$$

A consequence of (8) is that the closed loop system satisfies $\|A + BK\|_\infty \leq \lambda^* < 1$. Such a system is called superstable [25] when $\lambda^* < 1$. Superstable systems satisfy the following decay relations with respect to θ^*:=$\|E\|_1/(1 - \lambda^*)$ [25, Theorem A.1]

$$\begin{aligned}
\|w\|_\infty = 0 &\implies \quad \|x_t\|_\infty \leq (\lambda^*)^t\|x_0\|_\infty && \text{(9)}\\
\|w\|_\infty \leq 1 &\implies \quad \|x_t\|_\infty \leq \theta^* + (\lambda^*)^t \max(\|x_0\|_\infty - \theta^*, 0).
\end{aligned}$$

2) $\mathcal{H}_2$ Control: We now consider the case where w is an impulse, and the goal is to minimize the energy of z (the $\mathcal{H}_2$ problem). This problem can be recast into the following LMI

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

form [26], [27]:

$$\begin{aligned}\gamma_*^2 = \min_{Y,Z,S}\ &\gamma^2 \text{ subject to} \\ &\begin{bmatrix} Y - EE^T & AY+BS \\ * & Y\end{bmatrix} \in \mathbb{S}_+^{2n} \\ &\begin{bmatrix} Z & CY+DS \\ * & Y\end{bmatrix} \in \mathbb{S}_+^{n+r} \\ &\gamma^2 - \mathrm{Tr}(Z) \geq 0 \\ &Y \in \mathbb{S}_{++}^n,\ Z \in \mathbb{S}_+^r,\ S \in \mathbb{R}^{m \times n}.\end{aligned} \tag{10}$$

When (10) is feasible, the optimal feedback is $K = SY^{-1}$, and the corresponding $\mathcal{H}_2$ gain of (3) is γ_*^2. In this case, the function $x^TY^{-1}x$ is a Lyapunov function for the closed-loop system $x_+ = (A+BK)x$.

D. SOS Preliminaries

We briefly review notation used in defining Weighted Sum of Squares (WSOS) constraints for the imposition of Polynomial Matrix Inequalities (PMIs) certifying that a symmetric-matrix-valued polynomial is PSD.

A BSA set is defined by a finite number of bounded degree polynomials $\{g_i(x)\}_{i=1}^{N_g}$ and $\{h_j(x)\}_{j=1}^{N_h}$:

$$\mathbb{K}:=\{x \in \mathbb{R}^n \mid g_i(x) \geq 0,\ h_j(x) = 0\}. \tag{11}$$

A matrix-valued polynomial $P(x) \in \mathbb{S}^s[x]$ is PSD ($P(x) \in \mathbb{S}_+^s[x]$) if $\forall x \in \mathbb{R}^n:\ P(x) \in \mathbb{S}_+^s$. A matrix-valued-polynomial is an SOS matrix ($P(x) \in \Sigma^s[x]$) if there exists a vector of polynomials $v(x) \in \mathbb{R}[x]^q$ and a *Gram* matrix $Q \in \mathbb{S}_+^{qs}$ for some $q \in \mathbb{N}$ with $P(x) = (v(x) \otimes I_s)^T Q (v(x) \otimes I_s)$. When $v(x)$ is the set of monomials in n variables from degrees $0..d$, the Gram matrix Q has dimension $qs = \binom{n+d}{d}s$.

The set $\Sigma^s[\mathbb{K}]$ of WSOS matrices over (11) is the class of matrices $P(x) \in \mathbb{S}^s[x]$ such that there exists multipliers $\sigma_0(x) \in \Sigma^s[x],\ \sigma_i(x) \in \Sigma^s[x],\ \phi_j \in \mathbb{S}^s[x]$ with [28]

$$P(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x). \tag{12}$$

The set $\mathbb{K}$ is *Archimedean* if there exists an $R > 0$ such that the polynomial $R - \|x\|_2^2$ satisfies $R - \|x\|_2^2 \in \Sigma^1[\mathbb{K}]$. Archimedeanness of a set $\mathbb{K}$ is a property of its representation in (11). Every PD-valued matrix polynomial $P(x)$ over an Archimedean $\mathbb{K}$ satisfies $P(x) - \varepsilon I_q \succeq 0$ for some $\varepsilon > 0$ (Theorem 2 of [28], Scherer Positivestellensatz). The set $\Sigma^s[\mathbb{K}]$ is a subset of the set of matrix-valued-polynomials in x that are PSD over $\mathbb{K}$. The Putinar Positivestellensatz is the restriction of the Scherer Psatz to $s = 1$ [29].

III. Full Program

This section will present SOS approaches towards recovering robust controllers K (according to the criteria laid out in Section II-C) applicable for all plants consistent with data in $\mathcal{D}$. In this section we set $w = 0$ to simplify explanation and notation while still preserving the $A\Delta x$ and $B\Delta u$ bilinearities. Section VI re-introduces process noise $w \neq 0$ and modifies the developed method accordingly.

A. Consistency Sets

The consistency set $\bar{\mathcal{P}}(A, B, \Delta x, \Delta u)$ of plants and noise values $(A, B, \Delta x, \Delta u) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{n \times T} \times \mathbb{R}^{m \times (T-1)}$ that are consistent with data $\mathcal{D}$ under a noise bound of ϵ is described by

$$\bar{\mathcal{P}} := \begin{Bmatrix} \Delta x_{t+1} = A\Delta x_t + B\Delta u_t + h_t^0 & \forall t = 1..T-1 \\ \|\Delta x_t\|_\infty \leq \epsilon_x & \forall t = 1..T \\ \|\Delta u_t\|_\infty \leq \epsilon_u & \forall t = 1..T-1 \end{Bmatrix}, \tag{13}$$

with an intermediate definition of the affine weights h_t^0 as

$$h_t^0 := \hat{x}_{t+1} - A\hat{x}_t - B\hat{u}_t \qquad \forall t = 1..T-1. \tag{14}$$

Remark 1: Multiple observations $\{\mathcal{D}_k\}_{k=1}^{N_d}$ of the same system may be combined together by BSA intersections to form $\bar{\mathcal{P}} = \cap_{k=1}^{N_d} \bar{\mathcal{P}}(\mathcal{D}_k)$.

The set of plants $\mathcal{P}(A, B)$ consistent with the data in $\mathcal{D}$ is the projection

$$\mathcal{P}(A, B) := \pi^{A,B} \bar{\mathcal{P}}(A, B, \Delta x, \Delta u). \tag{15}$$

Remark 2: The consistency sets $\bar{\mathcal{P}}$ and $\mathcal{P}$ may be nonconvex and could even be disconnected, in a similar manner to the bilinear EIV representation in [15].

Remark 3: The describing constraints of $\bar{\mathcal{P}}$ are bilinear in terms of the groups (A, B) and $(\Delta x, \Delta u)$. Checking membership for fixed plant $(A_0, B_0) \in \mathcal{P}$ may be accomplished by solving a feasibility LP in terms of $(\Delta x, \Delta u)$.

Problem 1: The data driven robust performance problem is to find a control law $u = Kx$ that minimizes the worst case, over all pairs (A, B) in the consistency set $\mathcal{P}$, of the bound (6a) (ℓ^∞ setting) or γ in (10) ($\mathcal{H}_2$ setting).

B. Function Programs

This section will pose Problem 1 as a set of polynomial optimization programs, one for each performance criteria. All programs will require the following Assumption (for later convergence) :

Assumption 1: The sets $\bar{\mathcal{P}}$ (and therefore $\mathcal{P}$) are compact.

Assumption 1 may be satisfied if sufficient data is collected.

1) ℓ^∞ Robust Performance: A static gain $K \in \mathbb{R}^{m \times n}$ that achieves robust performance will be found by enforcing equation (6) for all plants in $\mathcal{P}$. The M matrix in (6) will be a matrix-valued function $M(A, B): \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \to \mathbb{R}^{n \times n}$. For a fixed $0 < \lambda < 1$, the problem reduces to

Problem 2:

$$\min_{v,\ S,\ M,\ \beta} \beta/(1-\lambda) \quad \text{subject to} \tag{16a}$$

$$\forall i \in 1..n:\ 1 \leq v_i \leq \beta \tag{16b}$$

$$\forall (A, B, \Delta x, \Delta u) \in \bar{\mathcal{P}}(A, B, \Delta x, \Delta u): \tag{16c}$$

$$\forall i = 1..n: \quad \lambda v_i - \textstyle\sum_{j=1}^n M_{ij}(A, B) \geq 0 \tag{16d}$$

$$\begin{aligned}&\forall i = 1..n,\ j = 1..n: \\ &\quad M_{ij}(A,B) - (A_{ij}v_j + \textstyle\sum_{\ell=1}^m B_{i\ell}S_{\ell j}) \geq 0 \\ &\quad M_{ij}(A,B) + (A_{ij}v_j + \textstyle\sum_{\ell=1}^m B_{i\ell}S_{\ell j}) \geq 0.\end{aligned} \tag{16e}$$

$$v \in \mathbb{R}_{>0}^n,\ \beta \in \mathbb{R},\ S \in \mathbb{R}^{m \times n}, \tag{16f}$$

$$M(A, B): \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m} \to \mathbb{R}^{n \times n}. \tag{16g}$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

A feasible solution (v, S, M) to Problem 2 could possess a discontinuous M function in (A, B). We now prove that if Problem 2 is feasible, then M can be chosen to be continuous (Lemma 3.1), and furthermore M can be chosen to be polynomial (Lemma 3.2).

Lemma 3.1: There exists a continuous selection for $M(A, B)$ given (v, S) under Assumption 1.

Proof: Define $\mathcal{M} : \mathcal{P} \rightrightarrows \mathbb{R}^{n\times n}_{+}$ as the set-valued map (solution region to (16d)-(16e)) described by

$$\mathcal{M}(A,B) := \left\{ \begin{array}{l} M \in \mathbb{R}^{n\times n}: \\ \quad \forall i: \ \sum_j M_{ij} \leq \lambda v_i \\ \quad \forall (i,j): \ -M_{ij} \leq \pm(A_{ij}v_j + \sum_{\ell=1}^m B_{i\ell}S_{\ell j}) \end{array} \right\}. \tag{17}$$

The right-hand sides of the constraints in (17) are each continuous (linear) functions of (A, B) as parameterized by (v, S). This linearity ensures that $\mathcal{M}$ is lower semi-continuous (Definition 1.4.2 of [30]) under the affine (continuous) changes in (A, B) in the compact domain $\mathcal{P}$ by Theorem 2.4 of [31] (perturbations of right-hand-sides of linear-inequality-defined regions). Michael's theorem (Proposition 9.3.2 in [30]) suffices to show that a continuous selection of $M \in \mathcal{M}(A, B)$ exists, given that $\mathcal{M}$ takes on closed convex values in the Banach space $\mathbb{R}^{n\times n}$, has a compact domain, and is lower-semicontinuous. One such continuous selection is the Minimal Map $M(A, B) := \operatorname{argmin}_{M' \in \mathcal{M}(\mathcal{A},\mathcal{B})} \|M'\|_F^2$. ■

Lemma 3.2: For any $(v, S, \tilde{M})$ solving Problem 2 with parameter $\lambda < 1$, (under Assumption 1), there exists a polynomial $M(A, B)$ such that (v, S, M) solves Problem 2 with parameter $\lambda' \in (\lambda, 1)$.

Proof:

Let $\epsilon > 0$ be a tolerance such that $\forall (i, j)$: $\sup_{(A,B)\in\mathcal{P}} |(\tilde{M}_{ij}(A, B) + \epsilon) - M_{ij}(A, B)| \leq \epsilon$ by the Stone-Weierstrass theorem in the compact set $\mathcal{P}$ [32]. This implies that $M(A, B) \geq \tilde{M}(A, B)$ everywhere in $\mathcal{P}$, because the residual $r_{ij}(A, B) = M_{ij}(A, B) - \tilde{M}_{ij}(A, B)$ takes on values between $[0, 2\epsilon]$. Now consider (16d) with M:

$$\begin{aligned} \lambda v_i - \textstyle\sum_{j=1}^n M_{ij} &= \lambda v_i - \textstyle\sum_{j=1}^n (\tilde{M}_{ij} + r_{ij}) && \text{(18a)} \\ &\geq \lambda v_i - 2\epsilon n - \textstyle\sum_{j=1}^n (\tilde{M}_{ij}). && \text{(18b)} \end{aligned}$$

For each row i, define Z_i^* as

$$Z_i^* := \sup_{(A,B)\in\mathcal{P}} \textstyle\sum_{j=1}^n (\tilde{M}_{ij}), \tag{19}$$

with the property that $\forall i : Z_i^* \leq \lambda v_i$. It therefore holds via (16d) that for some $\delta \in (0, 1 - \lambda)$:

$$\forall i = 1..n : \quad \lambda v_i - Z_i^* \geq 0 \quad \implies \quad (\lambda + \delta)v_i - Z_i^* > 0. \tag{20}$$

Substituting (20) into (18b) under the condition that (18b) must be nonnegative yields

$$((\lambda + \delta)v_i - Z_i^*) - 2\epsilon n \geq 0. \tag{21}$$

Choosing $\delta = 2\epsilon n/(\min_i v_i)$ and $\epsilon < \lambda(\min_i v_i)/(2n)$ (to ensure that $\lambda + \delta < 1$) will certify that M satisfies all inequality constraints w.r.t. $\lambda+\delta$, such as $(\lambda+\delta)v_i - \sum_{j=1}^n M_{ij} \geq 0$. The new parameter is $\lambda' = \lambda + \delta$. ■

Remark 4: A similar continuity analysis can take place over the $\bar{\mathcal{P}}$-worst-case superstabilization program derived from (8).

2) Robust $\mathcal{H}_2$ Control: . A static gain K that minimizes the worst case $\mathcal{H}_2$ norm over all plants $(A, B) \in \mathcal{P}$ can be found by solving:

Problem 3:

$$\begin{aligned} \gamma_0^2 = \min_{\gamma\in\mathbb{R}^+, Y, Z, S} \ & \gamma^2 \text{ subject to} && (22) \\ & \begin{bmatrix} Y - EE^T & AY + BS \\ * & Y \end{bmatrix} \in \mathbb{S}_+^{2n} && \forall (A, B, \Delta x, \Delta u) \in \bar{\mathcal{P}} \\ & \begin{bmatrix} Z & CY + DS \\ * & Y \end{bmatrix} \in \mathbb{S}_+^{n+r} \\ & \gamma^2 - \mathrm{Tr}(Z) \geq 0 \\ & Y \in \mathbb{S}_{++}^n, \ Z \in \mathbb{S}_+^r, \ S \in \mathbb{R}^{m\times n}. \end{aligned}$$

The corresponding feedback gain (independent of $(A, B, \Delta x, \Delta u)$) is given by $K = SY^{-1}$.
When the problem above is feasible, the function $x^T Y^{-1} x$ is a common Lyapunov function for all plants in $\mathcal{P}$. Hence the set is quadratically stabilizable [33]. Indeed, if one is only interested in quadratic stabilization, a simpler program can be obtained by setting $E = 0$ and enforcing only the first constraint in (22).

C. SOS Program and Numerical Considerations

Problems 2 and 3 may each be approximated by WSOS polynomials.

1) ℓ^∞ performance: SOS methods may be used to approximate (16) by requiring that $M(A, B) \in (\mathbb{R}[A, B])^{n\times n}$ is a polynomial matrix of degree $2d$. We note that $\bar{\mathcal{P}}$ from (15) is a BSA set, and will add the following assumption to ensure convergence:

Assumption 2: The sets $\bar{\mathcal{P}}$ and $\mathcal{P}$ are Archimedean.

Remark 5: While Assumption 2 is stronger than the compactness assumption 1, it holds in the practically relevant scenario where (loose) upper bounds on $\|A\|_F$ and $\|B\|_F$ are known. (These norm-bounding constraints are then added to the description of $\bar{\mathcal{P}}$ or $\mathcal{P}$).

Define $q_i^{\text{row}}(A, B, \Delta x, \Delta u; v, S, \lambda)$ as the left hand side of (16d), and let $q_{ij}^{\pm}(A, B, \Delta x, \Delta u; v, S)$ be the left hand side of each constraint in (16e). An example constraint from (16e) at (i, j) may be written as

$$\begin{aligned} & q_{ij}^-(A, B, \Delta x, \Delta u; v, S) \\ & = M_{ij}(A, B) - (A_{ij}v_j + \textstyle\sum_{\ell=1}^m B_{i\ell}S_{\ell j}). \end{aligned} \tag{23}$$

Equation (24) expresses the degree-$2d$ WSOS tightening of Problem 2, returning a controller $K = S\mathrm{diag}(1./v)$ if feasible:

$$\begin{aligned} J_d(\lambda) = \min_{v,S,M,\beta} \ & \beta/(1 - \lambda) \quad \text{subject to} \\ & 1 \leq v_i \leq \beta, \ S \in \mathbb{R}^{n\times m} && \text{(24a)} \\ & M \in \mathbb{R}[A, B]_{\leq 2d} && \text{(24b)} \\ & q_i^{row} \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} && \forall i \in 1..n \quad \text{(24c)} \\ & q_{ij}^{\pm} \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} && \forall i, j \in 1..n. \quad \text{(24d)} \end{aligned}$$

There are $2n^2 + n$ nonnegativity constraints in Program (24), each requiring a degree-$2d$ WSOS Psatz of (12). Each Psatz involves $n(n+T) + m(n+T-1)$ variables $(A, B, \Delta x, \Delta u)$, which induces a Gram matrices of maximal size $\binom{n(n+T)+m(n+T-1)+d}{d}$ at degree d.

Theorem 3.1: For every $\lambda \in [0, 1)$ and under Assumption 2, the objectives of program (24) will satisfy $J_d(\lambda) \geq J_{d+1}(\lambda) \geq \ldots$ and $\lim_{d\to\infty} J_d(\lambda) = J(\lambda)$.

Proof: This theorem follows from results on convergence of POPs. Applying the SOS hierarchy to a POP $p^* = \min_{x\in\mathbb{K}} p(x)$ for $\mathbb{K}$ Archimedean will result in a convergent sequence of lower bounds $p^*_d \leq p^*_{d+1} \ldots$ with $\lim_{d\to\infty} p^*_d = p^*$. The POP realization of nonnegativity program (16) is a feasibility problem with $p(x) = p^* = 0$. The lower bounds of the SOS hierarchy will therefore take a value of $p^*_d = 0$ (feasible, superstabilizing K found) or $p^*_d = -\infty$ (dual infeasible, superstabilizing K not found). By the limit property of $\lim_{d\to\infty} p^*_d = p^*$ under the Archimedean assumption 2, a superstabilizing K will be found if possible as the degree d increases. ■

2) $\mathcal{H}_2$ Control: The degree-d SOS truncation for (22) is

$$\begin{aligned}
\gamma^*_d = \min_{\gamma\in\mathbb{R}^+, Y, Z, S} \quad & \gamma^2 && (25)\\
& \begin{bmatrix} Y - EE^T & AY + BS \\ * & Y \end{bmatrix} \in \Sigma^{2n}[\bar{\mathcal{P}}]_{\leq 2d} \\
& \begin{bmatrix} Z & CY + DS \\ * & Y \end{bmatrix} \in \mathbb{S}^{n+r}_{+} \\
& \gamma^2 - \text{Tr}(Z) \geq 0 \\
& Y \in \mathbb{S}^n_{++}, Z \in \mathbb{S}^r_+, \ S \in \mathbb{R}^{m\times n}.
\end{aligned}$$

Problem (25) involves a PMI constraint for a matrix of size $2n$ and a PD constraint of size n. The PMI contains $n(n + T) + m(n + T - 1)$ variables $(A, B, \Delta x, \Delta u)$. The maximal Gram matrix size as induced by the degree-d Scherer Psatz (12) is $2n\binom{n(n+T)+m(n+T-1)+d}{d}$.

Theorem 3.2: Program (25) will return bounds $\gamma^*_d \geq \gamma^*_{d+1} \geq \ldots$ with convergence as $\lim_{d\to\infty} \gamma^*_d = \gamma^*_0$ from (22) Assumption 2.

Proof: This follows from the proof of Theorem 3.1 as modified for the Scherer Psatz in [28]. ■

The case where only quadratic stabilization is desired leads to the SOS program

$$\begin{aligned}
\text{find}_{Y,S} \quad & \begin{bmatrix} Y & AY + BS \\ * & Y \end{bmatrix} \in \Sigma^{2n}[\bar{\mathcal{P}}]_{\leq 2d} \\
& Y \in \mathbb{S}^n_{++}, \ S \in \mathbb{R}^{m\times n}. && (26)
\end{aligned}$$

Problem (26) will recover a Quadratically stabilizing K (if one exists) as $d \to \infty$ under Assumption 2. It has the same maximal Gram matrix size and convergence properties (Theorem 3.2) as the robust $\mathcal{H}_2$ case in (25).

IV. Alternatives Program

This section will formulate and use a Matrix Theorem of Alternatives in order to reduce the computational expense of running Algorithms (24) and (25). The cost savings are derived from elimination of the affine-entering noise variables $(\Delta x, \Delta u)$.

A. Theorem of Alternatives

Let $q : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \to \mathbb{S}^s$ be a symmetric-matrix valued function satisfying the constraint

$$q(A, B) \in \mathbb{S}^s_{++} \qquad \forall (A, B) \in \mathcal{P}. \tag{27}$$

If constraint (27) is satisfied (feasible), then the following problem is infeasible:

$$\text{find}_{(A,B)\in\mathcal{P}} \ -\lambda_{\min}(q(A, B)) \geq 0. \tag{28}$$

Lemma 4.1: Constraints (27) and (28) are strong alternatives (either one or the other is feasible).

Proof: If (27) is feasible, then $\lambda_{\min}(q(A, B))$ is positive for all $(A, B) \in \mathcal{P}$. Therefore, there does not exist an (A, B) such that $\lambda_{\min}(q(A, B)) \leq 0$, which is the statement of (28). On the opposite side, feasibility of (28) with an (A', B') : $\lambda_{\min}(q(A, B)) \leq 0$ implies that $\exists(A, B) \in \mathcal{P} : \ q(A, B) \notin \mathbb{S}^s_{++}$, and therefore (27) is infeasible. Additionally, there is no case where (27) and (28) are both infeasible: either all $q(A, B)$ are PD (27), or there exists a non-PD counterexample (28). ■

A set of dual variable functions $(\mu(A, B), \zeta(A, B), \psi(A, B))$ may be defined based on the constraint description from (13):

$$\begin{aligned}
\mu_{ti} &: \mathcal{P} \to \mathbb{S}^s && \forall i = 1..n, \ t = 1..T - 1 && (29a)\\
\zeta^{\pm}_{ti} &: \mathcal{P} \to \mathbb{S}^s_+ && \forall i = 1..n, \ t = 1..T && (29b)\\
\psi^{\pm}_{ti} &: \mathcal{P} \to \mathbb{S}^s_+ && \forall i = 1..n, \ t = 1..T - 1. && (29c)
\end{aligned}$$

The multipliers from (29), the Scherer Psatz (12), and the Robust Counterpart method of [34] can together be used to form the weighted sum $\Phi(A, B; \zeta^\pm, \psi^\pm, \mu) : \mathcal{P} \to \mathbb{S}^s$ with

$$\begin{aligned}
\Phi = &-q(A, B) && (30)\\
&+ \textstyle\sum_{t=1,i=1}^{T,n} \left((\epsilon_x - \Delta x_{ti})\zeta^+_{ti} + (\epsilon_x + \Delta x_{ti})\zeta^-_{ti}\right) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,m} \left((\epsilon_u - \Delta u_{ti})\psi^+_{ti} + (\epsilon_u + \Delta u_{ti})\psi^-_{ti}\right) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,n} (-\Delta x_{t+1,i} + A_i\Delta x_{ti} + B_i\Delta u_{ti} + h^0_{ti})\mu_{ti}.
\end{aligned}$$

The dual multipliers will always be treated as (possibly nonunique and discontinuous) functions $\mu(A, B)$, $\psi(A, B)$, or $\zeta(A, B)$, but their respective (A, B) dependence may be omitted to condense notation.

Theorem 4.1: A sufficient condition for infeasibility of program (28) is if there exists multipliers $\zeta^\pm, \psi^\pm, \mu$ according to (29) such that

$$\forall (A, B) \in \mathcal{P} : \sup_{\substack{\Delta x \in \mathbb{R}^{n\times T} \\ \Delta u \in \mathbb{R}^{m\times(T-1)}}} \lambda_{\max}(\Phi(A, B; \zeta^\pm, \psi^\pm, \mu)) < 0. \tag{31}$$

Proof: This theorem holds by arguments from [34], [35] with modifications for the matrix case.

Any point $(A, B, \Delta x, \Delta u) \in \bar{\mathcal{P}}$ must satisfy $\|\Delta x_t\|_\infty \leq \epsilon_x$ for all times $t \in 1..T$ and $\|\Delta u_t\|_\infty \leq \epsilon_x$ in times $t \in 1..T-1$ (13). This implies that $\epsilon_x \pm \Delta x_t$ and $\epsilon_u \pm \Delta u_t$ are nonnegative vectors for each time t for $(\Delta x, \Delta u) \in \pi^{\Delta x, \Delta u}\bar{\mathcal{P}}$, and therefore $(\epsilon_x \pm \Delta x_{ti})\zeta^{\pm}_{ti}$ and $(\epsilon_u \pm \Delta x_{ti})\psi^{\pm}_{ti}$ are each PSD matrices given that $\zeta^{\pm}_{ti}$ and $\psi^{\pm}_{ti}$ are PSD. Additionally, the data consistency constraints in $\Delta x_{t+1} + A\Delta x_t + B\Delta u_t + h^0_t$ from (13) evaluate to 0 for $(A, B, \Delta x, \Delta u) \in \bar{\mathcal{P}}$, so μ times these zero quantities result in a zero matrix. The addition of these PSD and Zero multiplier terms to $-q$ ensures that $\lambda_{\max}(\Phi) \geq \lambda_{\max}(-q)$. Finding multipliers $(\zeta^{\pm}, \psi^{\pm}, \mu)$ such that (31) holds therefore implies that $-q$ is Negative Definite (q is PD) over the space $\mathcal{P}$. This definiteness statement certifies infeasibility of (28), because there cannot exist a negative eigenvalue of q as (A, B) ranges over $\mathcal{P}$. ■

Theorem 4.2: The sufficient condition in Theorem 4.1 is additionally necessary for infeasibility of (28) in the case where $s = 1$ ($q(A, B)$ is scalar).

Proof: The term $\lambda_{\min}(q)$ is replaced by q in the scalar case of (28). The constraints in (13) are affine in $(\Delta x, \Delta u)$, which means they are both convex and concave in the variables $(\Delta x, \Delta u)$. The term $q(A, B)$ is also independent of $(\Delta x, \Delta u)$. Concavity of the constraints in $(\Delta x, \Delta u)$ is enough to establish necessity and strong alternatives by convex duality (Section 5.8 of [35]). ■

Remark 6: All of the inequalities in the extended superstabilizing program (16) have $s = 1$.

Remark 7: Refer to [36] for an example where the Alternatives procedure (31) with $s > 1$ is sufficient but not necessary (robust SDPs over interval matrices: "computationally tractable conservative approximation"). The work in [37] formulates robust SDPs involving polytopic uncertainty as a generally intractable two-stage optimization program.

Equation (31) can be simplified and transformed into a feasibility program by explicitly defining and constraining its $(\Delta x, \Delta u)$-supremal value. The sum Φ is affine in $(\Delta x, \Delta u)$, and the constant terms (in Δx and Δu) of Φ are arranged into $Q(A, B; \zeta^{\pm}, \psi^{\pm}, \mu)$:

$$\begin{aligned} Q := &-q(A,B) + \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x \left(\zeta^+_{ti} + \zeta^-_{ti}\right) \\ &+ \textstyle\sum_{t=1,i=1}^{T-1,m} \epsilon_u \left(\psi^+_{ti} + \psi^-_{ti}\right) + \textstyle\sum_{t=1,i=1}^{T-1,n} \mu_{ti} h^0_{ti}. \end{aligned} \tag{32}$$

The sum Φ may therefore be expressed as

$$\begin{aligned} \Phi = &Q + \textstyle\sum_{t=1,j=1,i=1}^{T-1,n,n} \mu_{tj}(A_{ji}\Delta x_{ti} + B_{ji}\Delta u_{ti}) \\ &+ \textstyle\sum_{t=1,i=1}^{T,n} (\zeta^-_{ti} - \zeta^+_{ti})\Delta x_{ti} + \textstyle\sum_{t=1,i=1}^{T-1,m} (\psi^-_{ti} - \psi^+_{ti})\Delta u_{ti} \\ &- \textstyle\sum_{t=2,i=1}^{T,n} \mu_{t-1,i}\Delta x_{ti} \end{aligned} \tag{33}$$

The supremal value of Φ in (31) given $(A, B; \zeta^{\pm}, \mu)$ is

$$\begin{aligned} \lambda^*(\Phi) &= \sup_{\Delta x, \Delta u} \lambda_{\max}(\Phi) \\ &= \begin{cases} \lambda_{\max}(Q) & \text{if } \zeta^+_{1i} - \zeta^-_{1i} = \sum_{j=1}^n A_{ji}\mu_{1j} \\ & \text{if } \zeta^+_{ti} - \zeta^-_{ti} = \sum_{j=1}^n A_{ji}\mu_{tj} - \mu_{t-1,i} \\ & \text{if } \zeta^+_{Ti} - \zeta^-_{Ti} = -\mu_{T-1,i} \\ & \text{if } \psi^+_{ti} - \psi^-_{ti} = \sum_{j=1}^m B_{ji}\mu_{tj} \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{34}$$

A feasibility program that ensures

$$\sup_{\Delta x, \Delta u} \lambda_{max}(\Phi(A, B)) < 0, \ \forall (A, B) \in \mathcal{P}, \ i = 1..n$$

is

$$\begin{aligned} &\text{find}_{\zeta,\mu} \ Q(A, B; \zeta^{\pm}, \psi^{\pm}, \mu) \prec 0 && \forall (A, B) \in \mathcal{P} && \text{(35a)} \\ &\zeta^+_{1i} - \zeta^-_{1i} = \textstyle\sum_{j=1}^n A_{ji}\mu_{1j} && && \text{(35b)} \\ &\zeta^+_{ti} - \zeta^-_{ti} = \textstyle\sum_{j=1}^n A_{ji}\mu_{tj} - \mu_{t-1,i} && \forall t \in 2..T-1 && \text{(35c)} \\ &\zeta^+_{Ti} - \zeta^-_{Ti} = -\mu_{T-1,i} && && \text{(35d)} \\ &\psi^+_{ti} - \psi^-_{ti} = \textstyle\sum_{j=1}^m B_{ji}\mu_{tj} && \forall t \in 1..T-1 && \text{(35e)} \\ &\zeta^{\pm}_{ti}(A, B) \in \mathbb{S}^s_+ && \forall t = 1..T, && \text{(35f)} \\ &\psi^{\pm}_{ti}(A, B) \in \mathbb{S}^s_+ && \forall t = 1..T-1, && \text{(35g)} \\ &\mu_{ti}(A, B) \in \mathbb{S}^s && \forall t = 1..T-1. && \text{(35h)} \end{aligned}$$

Theorem 4.3: When $s = 1$, program (35) is equivalent to the statement in (27), and is a strong alternative to (28) (exactly one of (27) or (28) are feasible). When $s > 1$, program (35) is a weak alternative to (28), and at most one of (28) and (35) are feasible.

Proof: Lemma 4.1 proves that (27) and (28) are strong alternatives.

In the case where $s = 1$, Theorem 4.2 proves that (28) and (31) are strong alternatives. Given that (35) is an explicit condition for validity of (31), it holds that (35) and (28) are strong alternatives and therefore (27) and (35) are equivalent.

In the more general case where $s > 1$, Theorem 4.1 proves that (28) and (31) are weak alternatives. Successfully finding a certificate $(\zeta^{\pm}, \psi^{\pm}, \mu)$ from (35) validates that (28) is infeasible. It is not possible for both (35) and (28) to hold simultaneously. ■

Theorem 4.4: Under Assumption 1, the dual multipliers $\zeta^{\pm}_{ti}(A, B)$ and $\mu_{ti}(A, B)$ which certify that $q(A, B) \succ 0$ over $\mathcal{P}$ via (35) may be chosen to be continuous, and may additionally be chosen to be polynomial.

Proof: See Appendix I for the proof of continuity, and Appendix II for the proof of polynomial selectability. ■

B. Alternatives SOS

The degree-$2d$ WSOS truncation of program (35) to certify positive definiteness in (27) is contained in Algorithm 1. The Alternatives Psatz in Algorithm 1 requires the following assumption to ensure convergence as $d \to \infty$:

Assumption 3: An Archimedean set $\Pi(A, B) \supseteq \mathcal{P}$ is previously known.

Remark 8: Such a set Π will describe the boundedness information on (A, B) discussed in Remark 5.

Remark 9: All of the equality constraints in (36e)-(36g) are linear constraints in the coefficients of $\zeta^{\pm}, \psi^{\pm}, \mu$.

The Full programs in (24) and (25) may be converted to Alternatives programs by replacing the respective cone containments $\in \Sigma^s[\bar{\mathcal{P}}]_{\leq 2d}$ by $\in \Sigma^{s,\text{alt}}[\Pi]_{\leq 2d}$. The below program (37) is an example of this type of Alternatives conversion for the ℓ_∞-regulation program (24) with the cone $\Sigma^{1,\text{alt}}[\Pi]_{\leq 2d}$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

Algorithm 1 Alternatives Psatz ($\Sigma^{s,\text{alt}}[\mathcal{P}]_{\leq 2d}$)

procedure ALTERN PSATZ($d,\ q(A,B), \Pi, \mathcal{D}, \epsilon, s$)

Solve (or find infeasibility certificate):

$$\zeta^{\pm}_{ti}(A,B) \in \Sigma^s[\Pi]_{\leq 2d} \tag{36a}$$
$$\psi^{\pm}_{ti}(A,B) \in \Sigma^s[\Pi]_{\leq 2d} \tag{36b}$$
$$\mu_{ti}(A,B) \in \mathbb{S}^s[A,B]_{\leq 2d-1} \tag{36c}$$
$$-Q(A,B;\zeta^{\pm},\mu) \in \Sigma[\Pi]_{\leq 2d} \text{ (from (32))} \tag{36d}$$
$$\zeta^{+}_{1i} - \zeta^{-}_{1i} = \textstyle\sum_{j=1}^{n} A_{ji}\mu_{1j} \tag{36e}$$
$$\zeta^{+}_{ti} - \zeta^{-}_{ti} = \textstyle\sum_{j=1}^{n} A_{ji}\mu_{tj} - \mu_{t-1,i} \quad \forall t \in 2..T-1 \tag{36f}$$
$$\zeta^{+}_{Ti} - \zeta^{-}_{Ti} = -\mu_{T-1,i} \tag{36g}$$
$$\psi^{+}_{ti} - \psi^{-}_{ti} = \textstyle\sum_{j=1}^{m} B_{ji}\mu_{tj} \quad \forall t \in 1..T-1. \tag{36h}$$

return $\zeta,\ \psi,\ \mu$ (or Infeasibility)
end procedure

rather than $\Sigma^1[\bar{\mathcal{P}}]_{\leq 2d}$:

$$\min \quad \beta/(1-\lambda) \tag{37a}$$
$$v \in [1,\beta]^n,\ \ S \in \mathbb{R}^{n\times m} \tag{37b}$$
$$M_{ij} \in \mathbb{R}[A,B]_{\leq 2d} \qquad \forall i,j \in 1..n \tag{37c}$$
$$q_i^{row} \in \Sigma^{1,\text{alt}}_{\leq 2d}[\Pi] \qquad \forall i \in 1..n \tag{37d}$$
$$q_{ij}^{\pm} \in \Sigma^{1,\text{alt}}_{\leq 2d}[\Pi] \qquad \forall i,j \in 1..n. \tag{37e}$$

Remark 10: Constraints (36a)- (36c) restrict $\zeta^{\pm}_{ti},\ \psi^{\pm}_{ti},\ \mu^{\pm}_{ti}$ to classes of bounded-degree polynomial matrices in (A,B) with degrees $\{2d, 2d, 2d-1\}$ respectively. Other approaches to truncation and verification infinite-dimensional linear programs can be used instead of polynomial/SOS methods, such as the neural-network based approach in [38] (with convergence in increasing number of neurons).

Remark 11: If Assumption 3 is violated (Π is a-priori unknown), then Equations (36) and (37) can be executed with $\zeta^{\pm}, \psi^{\pm} \in \Sigma^s[A,B]_{\leq 2d}$ ($\Pi = \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$). Program (37) can return sufficient conditions for ℓ_∞ regulation given λ, but it is no longer guaranteed that a λ-regulating controller will be recovered if feasible as $d \to \infty$.

V. Computational Complexity

This section tabulates the computational complexity involved in executing (Full or Alternative) WSOS tightenings of the robust performance programs 2 and 3 under measurement noise. As a reminder, the Full methods involve $p_F = n(n+T) + m(n+T-1)$ variables $(A, B, \Delta x, \Delta u)$, while the Alternatives methods employ $p_A = n(n+m)$ variables (A, B) after eliminating Δx and Δu.

The following tables will use the notation $\mathbb{R}[\cdot]$ for the size (cone) of a vector $\mathbb{R}^{(\cdot)\times 1}$ (free variable) and $\mathbb{S}_+[\cdot]$ for the size of a matrix $\mathbb{S}_+^{(\cdot)}$ (semidefinite variable).

The extended ℓ^∞-regulation algorithms (24) and (37) involve $2n^2+n$ polynomials $q(A,B)$ that must be nonnegative.

Table I compares the variable sizes in the Putinar Psatz (12) and the Alternatives Psatz (36).

TABLE I: Size of ℓ^∞-regulation Method (Psatz)

	Num. Polynomials	Full	Alternatives
q	1	$\mathbb{R}\left[\binom{p_F+2d}{2d}\right]$	$\mathbb{R}\left[\binom{p_A+2d}{2d}\right]$
σ_0	1	$\mathbb{S}_+\left[\binom{p_F+d}{d}\right]$	$\mathbb{S}_+\left[\binom{p_A+d}{d}\right]$
σ_i	$2nT$	$\mathbb{S}_+\left[\binom{p_F+d-1}{d-1}\right]$	$\mathbb{S}_+\left[\binom{p_A+d}{d}\right]$
μ_j	$2n(T-1)$	$\mathbb{R}\left[\binom{p_F+2d-2}{2d-2}\right]$	$\mathbb{R}\left[\binom{p_A+2d-1}{2d-1}\right]$

The $\mathcal{H}_2$-regulation algorithm (25) involves a single PMI of size $2n$, along with matrices $Y \in \mathbb{S}^{2n}_{++}$ and $Z \in \mathbb{S}^{r}_{+}$. Let $\nu = 2n(2n+1)/2$ be the number of free variables in a symmetric matrix in $\mathbb{S}^{2n}$. Table II lists the size of the matrices involved in the quadratically stabilizing Scherer Psatz expressions (12) and (36).

TABLE II: Size of $\mathcal{H}_2$-regulation Method (Psatz)

	Num. Polynomials	Full	Alternatives
q	1	$\mathbb{R}\left[\nu\binom{p_F+2d}{2d}\right]$	$\mathbb{R}\left[\nu\binom{p_A+2d}{2d}\right]$
σ_0	1	$\mathbb{S}_+\left[2n\binom{p_F+d}{d}\right]$	$\mathbb{S}_+\left[2n\binom{p_A+d}{d}\right]$
σ_i	$2nT$	$\mathbb{S}_+\left[2n\binom{p_F+d-1}{d-1}\right]$	$\mathbb{S}_+\left[2n\binom{p_A+d}{d}\right]$
μ_j	$2n(T-1)$	$\mathbb{R}\left[\nu\binom{p_F+2d-2}{2d-2}\right]$	$\mathbb{R}\left[\nu\binom{p_A+2d-1}{2d-1}\right]$

The main reduction in computational complexity stems from the fact that the alternatives Gram matrix sizes are independent from T. Table III displays the size (but not multiplicities) of variables involved in the ℓ^∞-regulation Algorithms (24), (37) for parameters of $n=2, m=1, d_{full}=2, d_{altern}=1$ and increasing T. Table IV lists sizes of SOS-tightenings of the $\mathcal{H}_2$-regulation Algorithm (26) for the same parameter values.

TABLE III: Size of Variables for ℓ^∞-regulation

	q	σ_0	σ_i	μ_j
Alternatives	28	7	7	7
Full(T = 4)	5985	171	18	171
Full(T = 6)	17550	300	24	300
Full(T = 8)	40920	465	30	465

TABLE IV: Size of Variables for $\mathcal{H}_2$-regulation

	q	σ_0	σ_i	μ_j
Alternatives	280	28	28	70
Full(T = 4)	59850	684	72	1710
Full(T = 6)	175500	1200	96	3000
Full(T = 8)	409200	1860	120	4650

VI. Including Process Noise

This section reinserts the process noise w from the model (1) into the description of plant consistency sets.

A. Set Description and Full

Let $(\epsilon_x, \epsilon_u, \epsilon_w) \geq 0$ be ℓ^∞ bounds for the measurement, input, and process noise respectively. The data $\mathcal{D} = (\tilde{x}_t, \tilde{u}_t)$ produces a consistency set $\bar{\mathcal{P}}^{\text{all}}$ as

$$\bar{\mathcal{P}}^{\text{all}}: \begin{Bmatrix} \|\Delta x_t\|_\infty & \leq \epsilon_x, & \forall t = 1..T \\ \|\Delta u_t\|_\infty & \leq \epsilon_u, & \forall t = 1..T-1 \\ \|w_t\|_\infty & \leq \epsilon_w, & \forall t = 1..T-1 \end{Bmatrix}, \tag{38}$$

with the defined quantities for all $t = 1..T-1$:

$$0 = -\Delta x_{t+1} + A\Delta x_t + B\Delta u_t + w_t + h_t^{\text{all}} \tag{39a}$$
$$h_t^{\text{all}} = \tilde{x}_{t+1} - A\tilde{x}_t - B\tilde{u}_t. \tag{39b}$$

The set $\bar{\mathcal{P}}^{\text{all}}$ in (38) is described by $2((T-1)(2n+m)+n)$ polynomial inequality constraints and $n(T-1)$ linear equality constraints in terms of the $p_F^{\text{all}} = n(n+m) + (T-1)(2n+m)+n$ variables $(A, B, \Delta x, \Delta u, w)$. Just as in (13) with $\mathcal{P}$ and $\bar{\mathcal{P}}$, the semialgebraic set of consistent plants $\mathcal{P}^{\text{all}}(A, B)$ may be formed by the projection

$$\mathcal{P}^{\text{all}}(A, B) = \pi^{A,B}\bar{\mathcal{P}}^{\text{all}}(A, B, \Delta x, \Delta u, w). \tag{40}$$

The combination of measurement, input, and process noise may be incorporated into the Full algorithms (24) (superstability) and (26) (quadratic stability) by imposing Psatz (12) positivity constraints over the set $\bar{\mathcal{P}}^{\text{all}}$ in (38).

Remark 12: We note that structured process noise $x_{k+1} = Ax_k + Bu_k + Ew_k$ for $E \in \mathbb{R}^{n\times e}$ may be incorporated into the All-noise framework. Additionally, the process noise variables w may be eliminated when $e \leq n$. Define a left inverse $E^\dagger$ with $E^\dagger E = I_e$ and a matrix $\mathcal{N}$ containing a basis for the nullspace of E^T in its columns. The following equations may be imposed to represent the process noise constraints for all time indices $t \in 1..T-1$:

$$\begin{aligned}\epsilon_w &\geq \|E^\dagger\left((\tilde{x}_{t+1} - \Delta x_{t+1}) - A(\tilde{x}_t - \Delta x_t) - B\tilde{u}_t\right)\|_\infty \\ 0 &= \mathcal{N}\left((\tilde{x}_{t+1} - \Delta x_{t+1}) - A(\tilde{x}_t - \Delta x_t) - B\tilde{u}_t\right).\end{aligned}$$

B. Alternatives

The BSA set (38) is described by $n(n+m)+T(2n+m)-n$ variables $(A, B, \Delta x, \Delta u, w)$. The variables $(\Delta x, \Delta u, w)$ may be eliminated by following the Theorem of Alternatives laid out in Section IV. The process noise variables w_t under the constraint $\forall t = 1..T-1, \ \|w_t\|_\infty \leq \epsilon_w$ will be eliminated by rearranging terms in (2):

$$\begin{aligned}w_t &= A(\hat{x}_t - \Delta x_t) + B(u_t - \Delta u_t) - (\hat{x}_{t+1} - \Delta x_{t+1}) \\ &= A\Delta x_t + B\Delta u_t + h_t^{\text{all}} - \Delta x_{t+1}.\end{aligned} \tag{41}$$

A certificate of PD-ness for the following matrix function $q(A, B)$ will be derived (from (27)):

$$q(A, B) \in \mathbb{S}_{++}^s \qquad \forall (A, B) \in \mathcal{P}^{\text{all}}. \tag{42}$$

Dual variables may be initialized with $\mu^\pm \in (\mathbb{S}_+^s[A, B])^{n\times(T-1)}$ for ϵ_w, $\psi^\pm \in (\mathbb{S}_+^s[A, B])^{n\times T}$ for ϵ_u, and $\zeta^\pm \in (\mathbb{S}_+^s[A, B])^{n\times T}$ for ϵ_x, according to the Putinar multipliers (ϕ, σ) from (12), to form the weighted sum Φ^{all} with

$$\begin{aligned}\Phi^{\text{all}} = -q(A, B) &+ \textstyle\sum_{t=1,i=1}^{T,n}\left(\zeta_{ti}^+(\epsilon_x - \Delta x_t) + \zeta_{ti}^-(\epsilon_x + \Delta x_t)\right) \\ &+ \textstyle\sum_{t,i}^{T,m}\left(\psi_{ti}^+(\epsilon_u - \Delta u_t) + \psi_{ti}^-(\epsilon_u + \Delta u_t)\right) \\ &+ \textstyle\sum_{t,i}^{T-1,n}\mu_{ti}^-(\epsilon_w - (A_i\Delta x_{ti} + B_i\Delta u_{ti} + h_{ti}^{\text{all}} - \Delta x_{t+1,i})) \\ &+ \textstyle\sum_{t,i}^{T-1,n}\mu_{ti}^+(\epsilon_w + (A_i\Delta x_{ti} + B_i\Delta u_{ti} + h_{ti}^{\text{all}} - \Delta x_{t+1,i})).\end{aligned} \tag{43}$$

The $(\Delta x, \Delta u)$-constant terms of Φ^{all} are

$$\begin{aligned}Q^{\text{all}} = -q(A, B) &+ \textstyle\sum_{t=1,i=1}^{T-1,n}\epsilon_w(\mu^- + \mu^+) \\ &+ \textstyle\sum_{t=1,i=1}^{T-1,n} h_{ti}^{\text{all}}(\mu^+ - \mu^-) \\ &+ \textstyle\sum_{t=1,i=1}^{T,m}\epsilon_u\left(\psi_{ti}^+ + \psi_{ti}^-\right) + \sum_{t=1,i=1}^{T,n}\epsilon_x\left(\zeta_{ti}^+ + \zeta_{ti}^-\right).\end{aligned} \tag{44}$$

Define the following symbol $\Delta\zeta = \zeta^+ - \zeta^-$, with a similar structure holding for $\Delta\psi$ and $\Delta\mu$. Following the supremum ≤ 0 procedure of Section IV leads to an alternatives-based nonnegativity certificate of (42):

$$\operatorname*{find}_{\zeta,\mu,\psi} \ -Q^{\text{all}}(A, B;\ \zeta^\pm, \mu^\pm, \psi^\pm) \in \mathbb{S}_+^s \qquad \forall (A, B) \tag{45a}$$
$$\Delta\zeta_{1i} = \textstyle\sum_{j=1}^n A_{ji}\left(\mu_{1j}^+ - \mu_{1j}^-\right) \tag{45b}$$
$$\Delta\psi_{ti} = \textstyle\sum_{j=1}^n B_{ji}\Delta\mu_{ti} \qquad \forall t = 1..T-1 \tag{45c}$$
$$\Delta\zeta_{ti} = \textstyle\sum_{j=1}^n A_{ji}\Delta\mu_{tj} - \Delta\mu_{t-1,i} \qquad \forall t \in 2..T-1 \tag{45d}$$
$$\Delta\zeta_{Ti} = -\Delta\mu_{T-1,i} \tag{45e}$$
$$\psi_{ti}^\pm,\ \zeta_{ti}^\pm \in \mathbb{S}_+^s[A, B] \qquad \forall i,\ \forall t \in 1..T \tag{45f}$$
$$\mu_{ti}^\pm \in \mathbb{S}_+^s[A, B] \qquad \forall i,\ \forall t \in 1..T-1. \tag{45g}$$

The certificate (45) involves only the $n(n+m)$ variables (A, B) at the cost of requiring $2T(2n+m)-2n+1$ Scherer Psatz constraints in (A, B) ((45a) and (45f)-(45g)). The cone of matrix-valued polynomials that admit certificates in (45) may be expressed as $\Sigma^s[A, B]^{\text{alt,all}}$. The cone $\Sigma^n[A, B]^{\text{alt,all}}_{\leq d}$ may be substituted in for $\Sigma^s[A, B]^{\text{alt}}_{\leq d}$ in all WSOS tightenings, such as in Algorithm (37). (superstability). An Archimedean set $\Pi^{\text{all}} \supseteq \mathcal{P}^{\text{all}}$ must be known to ensure convergence of the Alternatives certificate as the degree $d \to \infty$ (from Assumption 1). All Scherer constraints would then take place under $-Q,\ \mu_{ti}^\pm,\ \zeta_{ti}^\pm\ \psi_{ti}^\pm \in \Sigma^s[\Pi^{\text{all}}]_{\leq 2d}$ at degree-d.

VII. Numerical Examples

In this section we illustrate the advantages of the proposed method with several examples. First, we show that attempting to explain data generated by an EIV model using only process noise can lead to a controller that fails to stabilize the true plant. Then, we discuss the performance of all three stabilization algorithms. This performance is empirically compared using Monte Carlo experiments by adjusting the noise level and number of samples of data in $\mathcal{D}$. Finally, we investigate the worst-case $\mathcal{H}_2$ optimal control. A followup numerical experiment shows the result where all types of noises are considered. Finally, we illustrate that partial plant information helps to identify a controller.

MATLAB (2021a) code to generate the examples is publicly available[1]. Dependencies include Mosek [39] and YALMIP [40].

[1] https://github.com/jarmill/error_in_variables

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

A. EiV vs. Process Noise Only

This example illustrates the inadequacy of using process noise models to design controllers from data generated in an EIV scenario. Assume that the true data is generated by the open loop unstable system

$$A = \begin{bmatrix} 0.8779 & -1.2450 \\ 0.3229 & 1.0599 \end{bmatrix} \; B = \begin{bmatrix} 0.4901 & -0.4777 \\ 1.1646 & -0.9336 \end{bmatrix} \tag{46}$$

with noise levels $\epsilon_x = 0.0232, \epsilon_u = 0.0133; , \epsilon_\omega = 0.0042$. The corresponding noisy trajectories are

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.1761 & -1.2377 & -1.8334 & -0.5104 & 1.1809 & 3.5760 & 5.6699 \\ 0.8883 & 0.3922 & -0.8258 & -1.4751 & -2.1782 & -2.2015 & -1.7807 \end{bmatrix}$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -0.1792 & -1.0158 & -0.6882 & -0.6198 & -0.1990 & -0.1809 \\ 0.4558 & -0.3985 & -0.8140 & -0.3163 & 0.0910 & 0.3804 \end{bmatrix}$$

To find the minimum noise level necessary to explain the above ground-truth data using a process noise model we solved the following LP

$$\begin{aligned} \epsilon_\kappa = \; & \min_{A,B,\epsilon} \epsilon \\ & \|\hat{x}_{t+1} - A\hat{x}_t - B\hat{u}_t\|_\infty \le \epsilon \qquad \forall t \in 1..T-1 \\ & \epsilon \in \mathbb{R}, \;\; A \in \mathbb{R}^{n\times n}, \;\; B \in \mathbb{R}^{n\times m}. \end{aligned} \tag{47}$$

using the trajectories above, yielding $\epsilon_\kappa = 0.0311$. Using the algorithm 1 to design a data-driven controller assuming a process noise model with this bound leads to the controller

$$K = \begin{bmatrix} 1.5515 & -4.0390 \\ 2.5919 & -4.5144 \end{bmatrix}$$

which fails to stabilize the ground truth plant (46).

B. ℓ^∞ Induced Norm Minimization

This example involves minimization of the ℓ^∞-induced norm over all systems in $\bar{\mathcal{P}}$. The ground-truth system for this example is

$$A = \begin{bmatrix} 0.6863 & 0.3968 \\ 0.3456 & 1.0388 \end{bmatrix} \qquad B = \begin{bmatrix} 0.4170 & 0.0001 \\ 0.7203 & 0.3023 \end{bmatrix}, \tag{48}$$

and $E = I_2$.

A trajectory of length $T = 12$ is collected under the noise bounds $(\epsilon_x = 0.04, \epsilon_u = 0.04, \epsilon_w = 0)$. The Alternatives WSOS program in (37) is solved at $d = 1$ with 100 equally spaced parameter choices of $\lambda \in \{0, 0.01, 0.02, \ldots, 0.99\}$. The minimal upper-bound for the ℓ^∞ gain over these sample points takes place at $\lambda = 0.79$, with an ℓ^∞ upper-bound 5.0172. The resultant vector v and controller K is

$$v = \begin{bmatrix} 1.0000 \\ 1.0536 \end{bmatrix}, \qquad K = \begin{bmatrix} -0.9123 & -0.9411 \\ 1.0485 & -0.9898 \end{bmatrix}. \tag{49}$$

The $d = 1$ Alternatives superstabilization problem over $\bar{\mathcal{P}}$ (with $v = \mathbf{1}_2$) results in a worst-case decay rate of $\lambda^* = 0.8013$ and an ℓ^∞ upper bound of $\theta^* = 5.0332$. While the bound of 5.0332 computed from superstability is larger than the bound 5.0172 based on extended superstability, the superstability bound 5.0332 required the solution of only a single SDP rather than iterating over all 100 λ sample values.

Figure 1 reports ℓ^∞ upper-bounds as a function of λ, in which the minimal sampled value is 5.0172. All sampled λ values below 0.7100 returned either infeasibility or a Mosek `UNKNOWN` status when Program (37) was evaluated at $d = 1$.

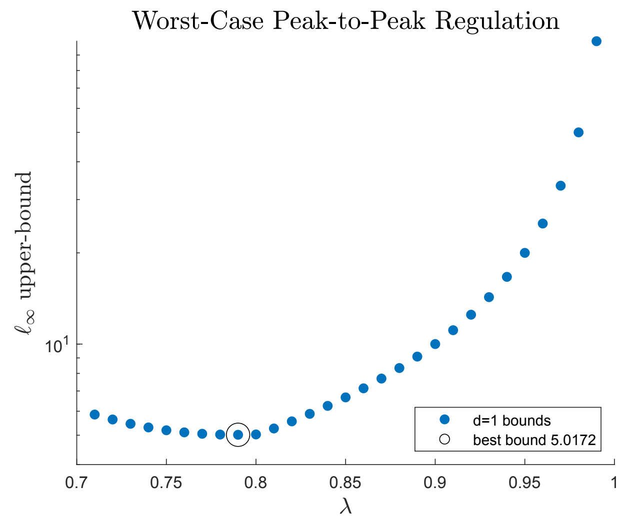


Fig. 1: ℓ^∞ bound vs. λ for peak-to-peak disturbance rejection

C. Monte Carlo Simulations for Stabilization

To test the reliability of the proposed method, we collected 50 trajectories with different level of noise and choose u, x_0 to be uniformly distributed in $[-1, 1]$. We then apply Alternatives WSOS tightenings of (16) and (22) on the following open-loop unstable system:

$$A^{\text{true}} = \begin{bmatrix} 0.6863 & 0.3968 \\ 0.3456 & 1.0388 \end{bmatrix}, \qquad B^{\text{true}} = \begin{bmatrix} 0.4170 & 0.0001 \\ 0.7203 & 0.3023 \end{bmatrix}. \tag{50}$$

We first investigate the effect of noise by fixing $T = 8$. The result is reported in TABLE V showing the percentage of successfully stabilized designs for extended superstability (ESS), superstability (SS), and quadratic stability (QS).

TABLE V: Success rate $(\%)$ as a function of ϵ with $T = 8$

ϵ	0.05	0.08	0.11	0.14
ESS	100	88	69	40
SS	100	84	57	39
QS	100	100	90	79

As expected, ESS performs better than SS, since ESS is a less restrictive stability condition. We note that QS was more successful than ESS in finding stabilizing controllers, but QS requires a maximal-size Gram matrix that is twice as large as in ESS. Increasing the noise level expands the consistency set, which in turn renders the problem of finding a single stabilizing controller more difficult. Collecting more sample data at the same noise bound $\epsilon = 0.14$ reduces the size of the

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2024.3447809

TABLE VI: Success rate (%) as a function of T with $\epsilon = 0.14$

T	8	10	12	14
ESS	40	61	78	89
SS	39	60	75	86
QS	79	86	95	99

consistency set, as illustrated for all stabilization methods in TABLE VI.

To see an advantage of QS over SS, we now consider a simple mass-spring-damper system shown below with $x_1 = x, x_2 = \dot{x}, u = F$ that includes integrator dynamics:

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} \tag{51}$$

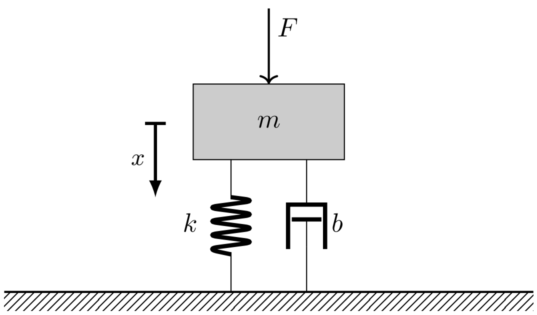


Fig. 2: Illustration of a Spring-Mass-Damper System with parameters in (51)

This system is not superstabilizable since any state feedback $u = Kx$ cannot affect the first row of A, hence the infinity norm of $A + BK$ is always greater or equal to 1. System (51) is extended-superstabilizable by [24, Example 4], and one can easily apply an Alternatives formulation (26) to try and find a quadratically stabilizing controller at the cost of multiplying the size of all Gram matrices involved by 2.

D. Worst-Case $\mathcal{H}_2$ Optimal Control

To analyze the $\mathcal{H}_2$ performance of the proposed method, we first solve the standard $\mathcal{H}_2$ problem (10) ($C = I_{r\times n},\ D = [\mathbf{0}_{n,m}; I_m],\ E = I_{n\times e}$), with known A, B defined by (50). We denote the benchmark as $\gamma_2 = 1.9084$. Now we apply Program (25) to 50 trajectories with different level of noise. The effect of noise is shown in TABLE VII with fixed $T = 8$.

For each of the 50 trajectories, Algorithm (25) returns a control policy K and a worst-case-$\mathcal{H}_2$ upper bound $\gamma_{2,worst}$ (valid for all $(A, B) \in \mathcal{P}$). The quantity $\gamma_{2,clp}$ (closed-loop poles) is the H_2 norm found by applying K as an input to the ground-truth system and solving (10). It therefore holds that $\gamma_{2,worst} \geq \gamma_{2,clp}$ for each trajectory. The quantities returned in Table VII and all subsequent tables are the median values of $(\gamma_{2,clp}, \gamma_{2,worst})$ over the 50 trajectories in order to prevent outliers deviating results.

As we increase the noise, $\gamma_{2,worst}$ also increases since the consistency set is expanded. However, $\gamma_{2,clp}$ does not necessarily increase since we only optimize in the worst case.

TABLE VII: $\mathcal{H}_2$ performance as a function of ϵ with $T = 8$

ϵ	0.05	0.08	0.11	0.14
$\gamma_{2,clp}$	1.9684	2.0715	2.1773	2.1456
$\gamma_{2,worst}$	2.3004	2.7308	3.2279	4.3137

It is also worth noting that $\gamma_2 \leq \gamma_{2,clp} \leq \gamma_{2,worst}$ and the equality holds only if there is no noise. H_2 performance can be improved by collecting more samples as shown in TABLE VIII.

TABLE VIII: $\mathcal{H}_2$ performance as a function of T with $\epsilon = 0.08$

T	8	10	12	14
$\gamma_{2,clp}$	2.0715	1.9637	1.9373	1.9321
$\gamma_{2,worst}$	2.7308	2.4160	2.2328	2.2014

E. All Noise

Consider the following set of noise bounds:

$$\begin{aligned} &\text{set } 1: \ \epsilon_x = 0.03, \epsilon_u = 0.00, \epsilon_w = 0.00 \\ &\text{set } 2: \ \epsilon_x = 0.00, \epsilon_u = 0.02, \epsilon_w = 0.00 \\ &\text{set } 3: \ \epsilon_x = 0.00, \epsilon_u = 0.00, \epsilon_w = 0.05 \\ &\text{set } 4: \ \epsilon_x = 0.03, \epsilon_u = 0.02, \epsilon_w = 0.00 \\ &\text{set } 5: \ \epsilon_x = 0.00, \epsilon_u = 0.02, \epsilon_w = 0.05 \\ &\text{set } 6: \ \epsilon_x = 0.03, \epsilon_u = 0.02, \epsilon_w = 0.05 \end{aligned} \tag{52}$$

For system (50), we collected a single data trajectory of length $T = 8$ starting from an initial state of $x_1 = [1; 0]$ with u uniformly distributed in $[-1, 1]^2$. The worst-case $\mathcal{H}_2$ norm for the all-noise bounds in (52) is collected in Table IX:

TABLE IX: $\mathcal{H}_2$ performance for different sets of noise

set	1	2	3	4	5	6
$\gamma_{2,clp}$	1.9340	1.9131	1.9750	1.9615	2.1249	2.1659
$\gamma_{2,worst}$	2.0681	1.9628	2.1294	2.1554	2.5029	2.5973

The $\mathcal{H}_2$ norm for the nominal plant is $\gamma_2 = 1.9084$. It is clear that adding more type of noise expands the consistency set, hence leading to a larger worst-case H_2 norm.

F. Partial Information

Partial information of the values of (A, B) can be easily incorporated into the proposed framework. Instead of treating all entries of A, B as unknown variables, we can suppose q entries of (A, B) are known. There are now $n(n+m) - q$ free variables defining the consistency set, producing a smaller Gram matrix of $\binom{n(n+m)-q+d}{d}$ as compared to $\binom{n(n+m)+d}{d}$. This size reduction ensures that it is theoretically and computationally easier to find stabilizing controllers K when partial information is known. For instance, if we assume that the first row of A is known and apply Program (25) with $T = 8$, $\epsilon = 0.08$, we get $\gamma_{2,clp} = 1.9568, \gamma_{2,worst} = 2.2566$ as compared to $\gamma_{2,clp} = 2.0715, \gamma_{2,worst} = 2.7308$ in the first column of TABLE VII.

VIII. Conclusion

This paper formulated data-driven state-feedback stabilization and robust performance problems for systems corrupted by ℓ^∞-bounded measurement, process, and input noise. (Full) WSOS programs for ℓ^∞ induced performance, superstabilizability, robust $\mathcal{H}_2$ and quadratic stabilizability will converge to their respective controllers (if such a controller exists) as the degree tends towards infinity. Such programs are computationally expensive with regards to the size of the PSD matrices required in SDPs. A theorem of alternatives was deployed to create equivalent (superstability) and conservative (quadratic stabilizability) programs at a reduced computational cost by eliminating the noise variables.

Future work involves developing MIMO dynamic-output-feedback controllers in the case of combined input noise and measurement noise. Other work involves analyzing conditions for which the sets $\bar{\mathcal{P}}$ and $\mathcal{P}$ are compact in terms of the collected data $\mathcal{D}$ (e.g. sampling complexity), and quantifying the conservativeness involved with utilizing the Theorem of Alternatives in the $s > 1$ case. Tractability of this method would improve with further development and reductions in complexity of solving SDPs.

Appendix I
Continuity of Multipliers

This appendix proves that the multiplier functions $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ from program (35) (when feasible) may be chosen to be continuous for any PD function $q(A,B)$ over $\mathcal{P}$ (Theorem 4.4). This proof will establish that a set-valued map based on a variant of the feasible set of (35) is lower semicontinuous, and will then apply Michael's theorem to certify continuous selections. The approach taken in this proof is similar to establishing lower semi-continuity of ρ-indexed sets from [41], but our problem has perturbations in the left-hand side multiplying the equality-multipliers μ.

We will pose the following assumption over the course of this appendix:

Assumption 4: The function $q(A,B) : \mathcal{P} \to \mathbb{S}^s_{++}$ is PD over the consistency set of plants $\mathcal{P}$.

The function $q(A,B)$ from Assumption 4 has a certificate of nonnegativity $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ by program (35). We note that $\mathcal{P}$ is compact by Assumption 1 and that the mapping from $\mathcal{P}$ to the constraints in (35) are affine (Lipschitz) in (A,B) over the compact $\mathcal{P}$.

Let $Z = ((\mathbb{S}^s)^{n\times T})^2 \times ((\mathbb{S}^s)^{m\times(T-1)})^2 \times (\mathbb{S}^s)^{n\times(T-1)}$ be the residing space (possible range) of the multipliers $(\zeta^+(A,B), \zeta^-(A,B), \psi^+(A,B), \psi^-(A,B), \mu(A,B))$. In this appendix, the notation $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ will refer to functions from (35), and variables $(\zeta^\pm, \psi^\pm, \mu)$ lacking arguments (A,B) will be values in Z.

A convex-set-valued map $S : \mathcal{P} \rightrightarrows Z$ may be defined as the feasible set of program (35) for each $(A,B) \in \mathcal{P}$. The domain of S is equal to $\mathcal{P}$ by Assumption 4 since the functions $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ have values in Z ($S(A,B) \neq \varnothing$) for all $(A,B) \in \mathcal{P}$. The range is nonclosed due to the PD constraint in (35a).

Define $\tau^* = \min_{(A,B)\in\mathcal{P}} \lambda_{\min}(q(A,B))$ as the minimal possible eigenvalue of the PD matrix $q(A,B)$. We note that the minimum is attained with $\tau^* > 0$ because $\mathcal{P}$ is compact and $\lambda_{\min}(q(A,B))$ is a continuous function of (A,B) (given that the eigenvalues of a matrix are continuous in the matrix entries).

We define the closed-convex-valued map $S^\tau : \mathcal{P} \rightrightarrows Z$ for a value $0 < \tau < \tau^*$ as the set of solutions $(\zeta^\pm, \psi^\pm, \mu) \in Z$ for each $(A,B) \in \mathcal{P}$ of

$$\begin{aligned}
&-Q(A,B;\zeta^\pm,\mu) - \tau I \in \mathbb{S}^s_+ && && \text{(53a)}\\
&\zeta^+_{1i} - \zeta^-_{1i} = \textstyle\sum_{j=1}^n A_{ji}\mu_{1j} && && \text{(53b)}\\
&\zeta^+_{ti} - \zeta^-_{ti} = \textstyle\sum_{j=1}^n A_{ji}\mu_{tj} - \mu_{t-1,i} && \forall t \in 2..T-1 && \text{(53c)}\\
&\zeta^+_{Ti} - \zeta^-_{Ti} = -\mu_{T-1,i} && && \text{(53d)}\\
&\zeta^\pm_{ti} \in \mathbb{S}^s_+, \ \mu_{ti} \in \mathbb{S}^s && \forall t = 1..T && \text{(53e)}\\
&\psi^+_{ti} - \psi^-_{ti} = \textstyle\sum_{j=1}^n B_{ji}\mu_{tj} - \mu_{t-1,i} && \forall t \in 1..T-1. && \text{(53f)}
\end{aligned}$$

The set-valued mappings S and S^τ are related by $S^\tau \subset S$ for all admissible τ.

We first note that Assumption 4 implies that $q(A,B) - \frac{\tau^*}{2}I \in \mathbb{S}^s_{++}$ over $(A,B) \in \mathcal{P}$. The set of multipliers that certify of PSD-ness for $q(A,B) - \frac{\tau^*}{2}I$ is $S^{\tau^*/2}$, which in turn is sufficient to prove that $q(A,B)$ is PD.

Any (possibly discontinuous) solution function $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ that certifies positive-definiteness of $q(A,B)$ over $\mathcal{P}$ via (35) satisfies

$$\begin{aligned}
&\forall (A,B) \in \mathcal{P}, \tau \in (0, \tau^*/2): \\
&\qquad (\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B)) \quad \in S^\tau(A,B). && \text{(54)}
\end{aligned}$$

A solution $(\zeta^\pm, \psi^\pm, \mu) \in S^\tau(A,B)$ is a *Slater point* if all matrices in (53a) and (53e) are PD and all equality constraints are fulfilled. A solution $(\zeta^\pm, \psi^\pm, \mu) \in S^{\tau'}(A,B)$ for some $\tau' > 0$ may be transformed into a new solution $(\tilde{\zeta}^\pm, \mu) \in S^{\tau/4}(A,B)$ such that $(\tilde{\zeta}^\pm, \tilde{\psi}^\pm, \mu)$ is a Slater point. Specifically, we express $-Q - \tau I$ from (32) as

$$\begin{aligned}
-Q - \tau I = q(A,B) &- \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x \left(\zeta^+_{ti} + \zeta^-_{ti} + \frac{\tau}{2Tn\epsilon_x} I\right) \\
&- \textstyle\sum_{t=1,i=1}^{T,m} \epsilon_x \left(\psi^+_{ti} + \psi^-_{ti} + \frac{\tau}{2Tn\epsilon_u} I\right) \\
&- \textstyle\sum_{t=1,i=1}^{T-1,n} \mu_{ti} h^0_{ti} - \frac{\tau}{2} I. && \text{(55a)}
\end{aligned}$$

The shifted multipliers $(\tilde{\zeta^\pm}, \psi^\pm)$ may be defined as

$$\tilde{\zeta}^\pm_{ti} = \zeta^\pm_{ti} + \frac{\tau}{8Tn\epsilon} I \qquad \tilde{\psi}^\pm_{ti} = \psi^\pm_{ti} + \frac{\tau}{8Tn\epsilon} I \qquad \text{(55b)}$$

which yields

$$\begin{aligned}
-Q - \tau I = q(A,B) &- \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x \left(\tilde{\zeta}^+_{ti} + \tilde{\zeta}^-_{ti}\right) \\
&- \textstyle\sum_{t=1,i=1}^{T-1,m} \epsilon_x \left(\tilde{\psi}^+_{ti} + \tilde{\zeta}^-_{ti}\right) \\
&- \textstyle\sum_{t=1,i=1}^{T-1,n} \mu_{ti} h^0_{ti} - \frac{\tau}{2} I && \text{(55c)}\\
= -\tilde{Q} - \tfrac{\tau}{2} I. && && \text{(55d)}
\end{aligned}$$

The differences in (53b)-(53d) cancel out with $\forall t, i:$

$$\tilde{\zeta}^+_{ti} - \tilde{\zeta}^-_{ti} = \left(\zeta^+_{ti} + \frac{\tau}{8Tn\epsilon}\right) - \left(\zeta^-_{ti} + \frac{\tau}{8Tn\epsilon}\right) = \zeta^+_{ti} - \zeta^-_{ti}, \qquad \text{(56)}$$

and likewise for $\psi^\pm$ and $\tilde{\psi}^\pm$

Lemma 1.1: The solution $(\tilde{\zeta}^\pm(A,B), \tilde{\psi}^\pm(A,B), \mu(A,B))$ constructed from a certificate $(\zeta(A,B), \psi(A,B), \mu(A,B))$ from (35) by (55b) is a Slater point of $S^{\frac{\tau}{8}}(A,B)$ for each $(A,B) \in \mathcal{P}$.

Proof: Given that $\zeta^\pm_{ti}(A,B)$ and $\psi^\pm_{ti}(A,B)$ are both PSD for each (t,i) (from (35f)), adding a PD matrix $\frac{\tau}{8Tn\epsilon}I$ to $\zeta^\pm_{ti}(A,B)$ will produce a PD $\tilde{\zeta}^\pm_{ti}(A,B)$ from (55b) (similar for ψ). The matrix $-\tilde{Q}$ from (55d) is also PD, since $-\tilde{Q} - \frac{\tau}{2}I \succeq 0 \implies -\tilde{Q} \succ \frac{\tau}{8}I$. All equality constraints remain feasible by the observation in (56), fulfilling the Slater point description. ∎

Lemma 1.2: The set-valued mapping S^τ is lower-semicontinuous over $\mathcal{P}$.

Proof: This follows from the (strong) Slater point characterization points in S^τ within $S^{\frac{\tau}{8}}$ from Lemma (1.1) by arguments from [42] extended to the Matrix case, given that S^τ has closed convex images and sends a compact set to a Banach space. ∎

The condition for Michael's theorem (Thm. 9.1.2 of [30]) to hold, guaranteeing a continuous selection of S^τ, is that $\mathcal{P}$ is compact, Z is a Banach space, S^τ is lower-semicontinuous, and S^τ has closed, nonempty, convex images for each $(A,B) \in \mathcal{P}$. All of these conditions hold, so a continuous selection $(\zeta^\pm_s, \psi^\pm_s, \mu_s) : \mathcal{P} \to Z$ may be chosen with $(\zeta^\pm_s(A,B), \psi^\pm_s(A,B), \mu_s(A,B)) \in S^\tau(A,B) \subset S(A,B)$. The continuous functions $(\zeta^\pm_s, \psi^\pm_s(A,B), \mu_s)$ may therefore be used to certify positive-definiteness of $q(A,B)$ over $\mathcal{P}$ in (35).

Appendix II
Polynomial Approximability of Multipliers

This appendix proves that a PD function $q(A,B)$ over $\mathcal{P}$ may be certified using polynomial multipliers $(\tilde{\zeta}^\pm(A,B), \tilde{\psi}^\pm(A,B), \tilde{\mu}(A,B))$ whenever (35) is feasible (Theorem 4.4). The proof will proceed through the introduction of three positive approximation tolerances $(\eta_0, \eta_1, \eta_2) > 0$ for use in the Stone-Weierstrass theorem over the compact set $\mathcal{P}$. For a matrix $F \in \mathbb{R}^{n\times m}$, define the element-wise maximum-absolute-value operator as $\mathbf{Mabs}(F) = \max_{(i,j)} |F_{ij}|$.

Let $(\zeta^\pm(A,B), \psi^\pm(A,B), \mu(A,B))$ be a continuous multiplier certificate from (35), in which continuity was proven by Appendix I. A symmetric polynomial multiplier matrix $\tilde{\mu}_{ti} \in \mathbb{S}^s[A,B]$ can be created for each (t,i) using the Stone-Weiesrtrass theorem:

$$\sup_{(A,B)\in\mathcal{P}} \mathbf{Mabs}(\tilde{\mu}_{ti}(A,B) - \mu_{ti}(A,B)) < \eta_0. \tag{57}$$

A. Multiplier Bound

This subsection will approximate the ζ and ψ multipliers by polynomials. The following explanation will only detail ζ multipliers, as the ψ derivation follows identical steps. Let $\gamma_{ti} \in \mathbb{S}[A,B]$ be the right-hand-sides of constraints (35b)-(35d) given $\tilde{\mu}$ with $(\forall i = 1..n)$:

$$\gamma_{1i} = \textstyle\sum_{j=1}^n A_{ji}\tilde{\mu}_{1j} \tag{58a}$$
$$\gamma_{ti} = \textstyle\sum_{j=1}^n A_{ji}\tilde{\mu}_{tj} - \tilde{\mu}_{t-1,i} \qquad \forall t \in 2..T-1 \tag{58b}$$
$$\gamma_{Ti} = -\tilde{\mu}_{T-1,i}. \tag{58c}$$

The equations $\zeta^+_{ti} - \zeta^-_{ti} = \gamma_{ti}$ from constraints (35b)-(35d) have solutions that can be parameterized by a set of continuous symmetric-matrix-valued functions $\phi_{ti}(A,B) : \mathcal{P} \to \mathbb{S}^s$:

$$\forall (t,i): \quad \zeta^+_{ti}(A,B) = \phi_{ti}(A,B)/2 + \gamma_{ti}(A,B)/2 \tag{59a}$$
$$\zeta^-_{ti}(A,B) = \phi_{ti}(A,B)/2 - \gamma_{ti}(A,B)/2. \tag{59b}$$

The functions ϕ_{ti} may be η_1-approximated by polynomials $\tilde{\phi}_{ti}(A,B) \in \mathbb{S}^s[A,B]$ for each (t,i) in the compact region $\mathcal{P}$:

$$\sup_{(A,B)\in\mathcal{P}} \mathbf{Mabs}(\tilde{\phi}_{ti}(A,B) - \phi_{ti}(A,B)) < \eta_1. \tag{60}$$

A tolerance $\eta_2 > 0$ is introduced to define the polynomial approximators $\tilde{\zeta}_{ti}$ for each (t,i):

$$\tilde{\zeta}^+_{ti}(A,B) = \tilde{\phi}_{ti}(A,B)/2 + \gamma_{ti}(A,B)/2 + (\eta_2/2)I \tag{61a}$$
$$\tilde{\zeta}^-_{ti}(A,B) = \tilde{\phi}_{ti}(A,B)/2 - \gamma_{ti}(A,B)/2 + (\eta_2/2)I. \tag{61b}$$

The approximators $\tilde{\zeta}$ are related to the original multipliers ζ for each (t,i) by

$$\tilde{\zeta}^\pm_{ti}(A,B) = \zeta^\pm_{ti}(A,B) + (\tilde{\phi}_{ti}(A,B) - \phi_{ti}(A,B))/2 + (\eta_2/2)I. \tag{62}$$

The approximant $\tilde{\zeta}$ must take on PSD values in order to ensure that it is a valid multiplier for (35f).

Lemma 2.1: Let $M \in \mathbb{S}^s_{++}$ and $R \in \mathbb{S}^s$ be matrices with $\mathbf{Mabs}(R) \leq \eta$ for some $\eta > 0$. A sufficient condition for $M + R \in \mathbb{S}^s_{++}$ for all R is that $M - \eta s I \in \mathbb{S}^s_{++}$.

Proof: The minimum eigenvalue of $M+R$ is lower-bounded by Weyl's inequality for the sum of Hermitian matrices as

$$\lambda_{\min}(M+R) \geq \lambda_{\min}(M) + \min_{R\in\mathbb{S}^s,\ R_{ij}\in[-\eta,\eta]} \lambda_{\min}(R).$$

The minimum eigenvalue of R is in turn lower-bounded by $-\eta s$ through the Gershgorin circle theorem, implying that

$$\lambda_{\min}(M+R) \geq \lambda_{\min}(M) - \eta s. \tag{63}$$

A sufficient condition for (63) to hold is if $M - \eta s I \in \mathbb{S}^s_{++}$. ∎

The matrix $(\eta_2/2)I - (\tilde{\phi}_{ti}(A,B) - \phi_{ti}(A,B))/2$ from (62) may be analyzed using Lemma 2.1 with $M = (\eta_2/2)I$ and $\eta = (\eta_1)/2$. Given that $\zeta^\pm_{ti} \in \mathbb{S}^s_+(A,B)$, a sufficient condition for $\tilde{\zeta}^\pm_{ti} \in \mathbb{S}^s_+(A,B)$ by Lemma 2.1 is

$$\eta_2 > \eta_1 s. \tag{64}$$

B. Certificate Bound

Let us express $\psi^{\pm}$ in terms of functions $\tilde{\phi}^u_{ti}, \tilde{\gamma}^u_{ti}$ in a similar manner to (61) as

$$\tilde{\psi}^{\pm}_{ti}(A,B) = \tilde{\phi}^u_{ti}(A,B)/2 \pm \gamma^u_{ti}(A,B)/2 + (\eta_2/2)I. \quad (65)$$

The $(\Delta x, \Delta u)$-constant term Q in (32) has a polynomial approximation (when substituting $\zeta^{\pm} \to \tilde{\zeta}^{\pm},\ \psi^{\pm} \to \tilde{\psi}^{\pm}\ \mu \to \tilde{\mu}$) of

$$\begin{aligned}
\tilde{Q} &= -q(A,B) + \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x \left(\tilde{\zeta}^+_{ti} + \tilde{\zeta}^-_{ti}\right) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,m} \epsilon_u \left(\tilde{\psi}^+_{ti} + \tilde{\psi}^-_{ti}\right) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,n} \tilde{\mu}_{ti} h^0_{ti} && (66a)\\
&= Q + \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x (\tilde{\phi}_{ti} - \phi_{ti} + \eta_2 I) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,m} \epsilon_u (\tilde{\phi}^u_{ti} - \phi^u_{ti} + \eta_2 I) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti}) h^0_{ti} && (66b)\\
&= Q + \eta_2(\epsilon_x T n + \epsilon_u (T-1)m) I + \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_x (\tilde{\phi}_{ti} - \phi_{ti}) \\
&+ \textstyle\sum_{t=1,i=1}^{T-1,m} \epsilon_u (\tilde{\phi}^u_{ti} - \phi^u_{ti}) + \sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti}) h^0_{ti}. && (66c)
\end{aligned}$$

The negative of (66) is

$$\begin{aligned}
-\tilde{Q} = &- Q - \eta_2(\epsilon_x T n + \epsilon_u (T-1)m) I \\
&- \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_u (\tilde{\phi}_{ti} - \phi_{ti}) - \sum_{t=1,i=1}^{T-1,m} \epsilon_u (\tilde{\phi}^u_{ti} - \phi^u_{ti}) \\
&- \textstyle\sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti}) h^0_{ti}. && (67)
\end{aligned}$$

Lemma 2.1 will be used to derive a sufficient condition on (η_0, η_1, η_2) such that $-\tilde{Q}$ is PD for all $(A,B) \in \mathcal{P}$ (satisfying condition (35a)). Define the smallest eigenvalue of $-Q$ as

$$\lambda^* = \min_{(A,B)\in\mathcal{P}} \lambda_{\min}(-Q(A,B)) > 0. \quad (68)$$

Because Q satisfies (35a), $-Q$ is PD over $\mathcal{P}$ and therefore $\lambda^* > 0$. Further define $\bar{h}^0$ and $\bar{H}$ using (14) as

$$\begin{aligned}
\bar{h}^0_{ti} &:= \max_{(A,B)\in\mathcal{P}} h^0_{ti}(A,B), \quad \forall (t,i) && (69)\\
\bar{H} &:= \textstyle\sum_{t=1,i=1}^{T-1,n} \bar{h}^0_{ti}. && (70)
\end{aligned}$$

The expression in (67) therefore satisfies

$$\begin{aligned}
-\tilde{Q} \succeq &(-Q - \eta_2(\epsilon_x T n + \epsilon_u (T-1)m) I) - \\
&- \textstyle\sum_{t=1,i=1}^{T,n} \epsilon_u (\tilde{\phi}_{ti} - \phi_{ti}) - \sum_{t=1,i=1}^{T-1,m} \epsilon_u (\tilde{\phi}^u_{ti} - \phi^u_{ti}) && (71)\\
&- \textstyle\sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti}) \bar{h}^0_{ti}, && (72)
\end{aligned}$$

and we further note that

$$- \textstyle\sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti}) \bar{h}^0_{ti} \geq -\bar{H} \left(\sum_{t=1,i=1}^{T-1,n} (\tilde{\mu}_{ti} - \mu_{ti})\right). \quad (73)$$

We can apply Lemma 2.1 towards the right-hand-side of (72) under the definitions of

$$\begin{aligned}
M &:= -Q - \eta_2(\epsilon_x T n + \epsilon_u (T-1)m)) I && (74a)\\
\eta &:= (T n \epsilon_x + (T-1)m\epsilon_u)\eta_1 + \bar{H}\eta_0 && (74b)
\end{aligned}$$

to form the condition

$$\begin{aligned}
\lambda_{\min}(-\tilde{Q}) &\geq \lambda_{\min}(M) - \eta s && (75)\\
&\geq \lambda^* - \eta_2 s(\epsilon_x T n + \epsilon_u (T-1)m) \\
&\quad - s(T n \epsilon_x + (T-1)m\epsilon_u)\eta_1 + \bar{H}\eta_0. && (76)
\end{aligned}$$

The combined conditions for admissible (η_0, η_1, η_2) such that $-\tilde{Q}$ is PD are

$$\begin{aligned}
&\eta_0, \eta_1, \eta_2 > 0, \qquad \eta_2 > \eta_1 s && (77a)\\
&\lambda^* > (\epsilon_x T n + \epsilon_u (T-1)m)s)\eta_2 \\
&\quad + (T n \epsilon_x + (T-1)m\epsilon_u)s\eta_1 + (\bar{H}s)\eta_0. && (77b)
\end{aligned}$$

Under the definition

$$\bar{\eta} := (2s+1)(\epsilon_x T n + \epsilon_u (T-1)m)), \quad (78)$$

one possible choice of $(\eta_0, \eta_1, \eta_2) > 0$ satisfying (77) is

$$\eta_0 = \frac{\lambda^*}{4\bar{H}s} \qquad \eta_1 = \frac{\lambda^*}{2\bar{\eta}} \qquad \eta_2 = \frac{\lambda^*(s+1)}{2\bar{\eta}}. \quad (79)$$

A polynomial multiplier certificate $(\zeta^{\pm}, \mu)$ will therefore always exist whenever (35) is satisfied.

Acknowledgements

The authors would like to thank Didier Henrion, Milan Korda, Victor Magron, Antonio Bellin, and the POP group at LAAS-CNRS for their discussions and suggestions for improving the original manuscript.

References

[1] Y. Cheng, M. Sznaier, and C. Lagoa, "Robust Superstabilizing Controller Design from Open-Loop Experimental Input/Output Data," *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 1337–1342, 2015, 17th IFAC Symposium on System Identification SYSID 2015.

[2] T. Dai and M. Sznaier, "Data Driven Robust Superstable Control of Switched Systems," *IFAC-PapersOnLine*, vol. 51, no. 25, pp. 402–408, 2018.

[3] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4753–4768, 2020.

[4] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, "From Noisy Data to Feedback Controllers: Nonconservative Design via a Matrix S-Lemma," *IEEE Transactions on Automatic Control*, 2020.

[5] A. Bisoffi, C. De Persis, and P. Tesi, "Data-driven control via Petersen's lemma," *arXiv preprint arXiv:2109.12175*, 2021.

[6] J. Berberich, C. W. Scherer, and F. Allgöwer, "Combining prior knowledge and data for robust controller design," *IEEE Transactions on Automatic Control*, 2022.

[7] M. Bianchi, S. Grammatico, and J. Cortés, "Data-driven stabilization of switched and constrained linear systems," *arXiv preprint arXiv:2208.11392*, 2022.

[8] T. Dai and M. Sznaier, "A convex optimization approach to synthesizing state feedback data-driven controllers for switched linear systems," *Automatica*, vol. 139, p. 110190, 2022.

[9] ——, "Data-driven quadratic stabilization and lqr control of lti systems," *Automatica*, vol. 153, p. 111041, 2023.

[10] B. Shafai, A. Moradmand, and M. Siami, "Data-Driven Positive Stabilization of Linear Systems," in *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*, vol. 1. IEEE, 2022, pp. 1031–1036.

[11] J. Miller, T. Dai, M. Sznaier, and B. Shafai, "Data-Driven Control of Positive Linear Systems using Linear Programming," in *Proc. IEEE Conf. Decis. Control 2023*, 2023, pp. 1588–1594.

[12] B. Polyak and M. Halpern, "Optimal design for discrete-time linear systems via new performance index," *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 2, pp. 129–152, 2001.

[13] V. Cerone, "Parameter bounds for ARMAX models from records with bounded errors in variables," *International Journal of Control*, vol. 57, no. 1, pp. 225–235, 1993.

[14] J. Norton, "Identification of parameter bounds for ARMAX models from records with bounded noise," *International Journal of control*, vol. 45, no. 2, pp. 375–390, 1987.

[15] V. Cerone, "Feasible parameter set for linear models with bounded errors in all variables," *Automatica*, vol. 29, no. 6, pp. 1551–1555, 1993.

[16] V. Cerone, D. Piga, and D. Regruto, "Set-Membership Error-in-Variables Identification Through Convex Relaxation Techniques," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 517–522, 2011.

[17] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2020.

[18] J. Coulson, H. van Waarde, and F. Dörfler, "Robust Fundamental Lemma for Data-driven Control," *arXiv preprint arXiv:2205.06636*, 2022.

[19] A. Bisoffi, L. Li, C. D. Persis, and N. Monshizadeh, "Controller synthesis for input-state data with measurement errors," 2024.

[20] L. Li, A. Bisoffi, C. D. Persis, and N. Monshizadeh, "Controller synthesis from noisy-input noisy-output data," 2024.

[21] J. Miller, T. Dai, and M. Sznaier, "Data-Driven Superstabilizing Control of Error-in-Variables Discrete-Time Linear Systems," in *Proc. IEEE Conf. Decis. Control 2022*. IEEE, 2022, pp. 4924–4929.

[22] J. B. Lasserre, *Moments, Positive Polynomials And Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.

[23] M. Dahleh and I. Díaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Prentice Hall, 1995.

[24] B. T. Polyak, "Extended superstability in control theory," *Automation and Remote Control*, vol. 65, no. 4, pp. 567–576, 2004.

[25] B. T. Polyak and P. S. Shcherbakov, "Superstable Linear Control Systems. I. Analysis," *Automation and Remote Control*, vol. 63, no. 8, pp. 1239–1254, 2002.

[26] R. J. Caverly and J. R. Forbes, "LMI Properties and Applications in Systems, Stability, and Control Theory," *arXiv preprint arXiv:1903.08599*, 2019.

[27] E. Feron, V. Balakrishnan, S. Boyd, and L. El Ghaoui, "Numerical methods for H2 related problems," in *1992 American control conference*. IEEE, 1992, pp. 2921–2922.

[28] C. W. Scherer and C. W. Hol, "Matrix Sum-of-Squares Relaxations for Robust Semi-Definite Programs," *Mathematical programming*, vol. 107, no. 1, pp. 189–211, 2006.

[29] M. Putinar, "Positive Polynomials on Compact Semi-algebraic Sets," *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.

[30] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*. Springer Science & Business Media, 2009.

[31] O. L. Mangasarian and T.-H. Shiau, "Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems," *SIAM Journal on Control and Optimization*, vol. 25, no. 3, pp. 583–595, 1987.

[32] M. H. Stone, "The generalized Weierstrass approximation theorem," *Mathematics Magazine*, vol. 21, no. 5, pp. 237–254, 1948.

[33] B. R. Barmish, "Necessary and sufficient conditions for quadratic stabilizability of an uncertain system," *Journal of Optimization theory and applications*, vol. 46, no. 4, pp. 399–408, 1985.

[34] A. Ben-Tal, D. Den Hertog, and J.-P. Vial, "Deriving robust counterparts of nonlinear uncertain inequalities," *Mathematical Programming*, vol. 149, no. 1-2, pp. 265–299, 2015.

[35] L. Boyd, Stephen and Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[36] A. Ben-Tal and A. Nemirovski, "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty," *SIAM Journal on Optimization*, vol. 12, no. 3, pp. 811–833, 2002.

[37] J. Zhen, F. J. de Ruiter, E. Roos, and D. den Hertog, "Robust optimization for models with uncertain second-order cone and semidefinite programming constraints," *INFORMS Journal on Computing*, vol. 34, no. 1, pp. 196–210, 2022.

[38] A. Abate, D. Ahmed, A. Edwards, M. Giacobbe, and A. Peruffo, "Fossil: a software tool for the formal synthesis of lyapunov functions and barrier certificates using neural networks," in *Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control*, 2021, pp. 1–11.

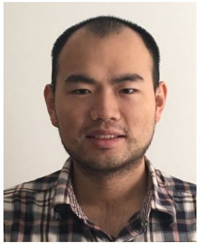
[39] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.*, 2020. [Online]. Available: https://docs.mosek.com/9.2/toolbox/index.html

[40] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[41] J. Denel, "Extensions of the continuity of point-to-set maps: Applications to fixed point algorithms," *Point-to-Set Maps and Mathematical Programming*, pp. 48–68, 1979.

[42] A. Daniilidis, M. A. Goberna, M. A. López, and R. Lucchetti, "Lower Semicontinuity of the Feasible Set Mapping of Linear Systems Relative to Their Domains," *Set-Valued and Variational Analysis*, vol. 21, no. 1, pp. 67–92, 2013.

Jared Miller is currently a Post-Doctoral Researcher at the Automatic Control Laboratory (IfA) at ETH Zürich, in the research group of Prof. Roy S. Smith. He received his B.S. and M.S. degrees in Electrical Engineering from Northeastern University in 2018, and his Ph.D. degree in Electrical Engineering from Northeastern University in 2023. He is a recipient of a 2020 Chateaubriand Fellowship from the Office for Science Technology of the Embassy of France in the United States. He was given an Outstanding Student Paper awards at the IEEE Conference on Decision and Control in 2021 and in 2022, and was a finalist for the Young Author Award at the 2023 IFAC World Congress. His research interests include large-scale convex optimization, renewable energy integration, measure theory, and analysis of nonlinear systems.



Tianyu Dai is currently a Senior Software Engineer in the Controls & Idnetification groups at MathWorks. He received his B.S. degree in Optical information of Science and Technology from Harbin Institute of Technology (Weihai), China in 2014. He received his Ph.D. degree in Electrical Engineering at Northeastern University, United States in 2022. His research interests lie in data-driven control, convex optimization and system identification.

Mario Sznaier is currently the Dennis Picard Chaired Professor at the Electrical and Computer Engineering Department, Northeastern University, Boston. Prior to joining Northeastern University, Dr. Sznaier was a Professor of Electrical Engineering at the Pennsylvania State University and also held visiting positions at the California Institute of Technology. His research interest include robust identification and control of hybrid systems, robust optimization, and dynamical vision. Dr. Sznaier is currently serving as Editor in Chief of the section on AI and Control of the journal Frontiers in Control Engineering, Vice-Chair of IFAC's CC2 Coordinating Committe, member of IFAC's Conference Board, and general chair of SysID 2024. Additional service includes Associate Editor for the Journal Automatica (2005-2021), chair of IFAC's Technical Committee TC 2.5.Technical Committee on Computational Aspects of Control System Design (2012-2016), General co-Chair and Program Chair of the 2017 CDC. He is a distinguished member of the IEEE Control Systems Society and an IEEE Fellow.