2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Data-Driven Superstabilizing Control of Error-in-Variables Discrete-Time Linear Systems

Jared Miller[1], Tianyu Dai [1], Mario Sznaier[1]

***Abstract*— This paper proposes a method to find superstabilizing controllers for discrete-time linear systems that are consistent with a set of corrupted observations. The L-infinity bounded measurement noise introduces a bilinearity between the unknown plant parameters and noise terms. A superstabilizing controller may be found by solving a feasibility problem involving a set of polynomial nonnegativity constraints in terms of the unknown plant parameters and noise terms. A sequence of sum-of-squares (SOS) programs in rising degree will yield a super-stabilizing controller if such a controller exists. Unfortunately, these SOS programs exhibit very poor scaling as the degree increases. A theorem of alternatives is employed to yield equivalent, convergent (under mild conditions), and more computationally tractable SOS programs.**

I. Introduction

The data-driven control problem has received renewed interest in the last few years, as an alternative to conventional approaches that first identify a model and then use it to design a controller. Given data generated by an (unknown) system of the form,

$$x_{t+1} = Ax_t + Bu_t + w_t \tag{1}$$

$$\hat{x}_t = x_t + \Delta x_t, \quad \hat{u}_t = u_t + \Delta u_t, \tag{2}$$

where $w_t,\ \Delta x_t,\ \Delta u_t$ respectively represent process, measurement, and input noise, the goal is to use measured data $(\hat{u}_t, \hat{x}_t)$ to find a gain K such that $(A + BK)$ is Hurwitz, for all possible pairs (A, B) consistent with this data.

For the case of noise-free data, the milestone paper [1] based on the Willem's fundamental lemma [2] parameterized the controller directly from the input/output data. The case of systems subject to process noise only (e.g. $\Delta x, \Delta u \equiv 0$) has been well studied, under different scenarios: [3]–[5] focus on control under bounded-energy (ℓ_2 norm) process noise, and solve this type of problem by polynomial-time Semidefinite Programming SDP. The work in [5] provides further discussion on the ℓ_∞ bounded process noise setting. This setting is more desirable than the ℓ_2 norm in many scenarios, since it allows for considering noise bounds that are independent of the measurement horizon [5]. Thus, data can be added as it becomes available during operation. In addition, these ℓ_∞ error bounds arise naturally when the discrete system (1) originates from the discretization of a continuous time sysetm, in which case w_t models the error when approximating the time derivative with finite differences. Finally, ℓ_∞ bounds are relevant when the goal is to design ℓ_1 optimal controllers capable of handling time-varying uncertainty. Unfortunately, the computational complexity of handling ℓ_∞ bounded uncertainty grows exponentially with the number of measurements [5]. A tractable alternative to handle ℓ_∞ noise was proposed in [6]–[8], based on the concept of superstability. Superstability is more conservative than stability, but it may be solved in a tractable manner through convex optimization and it also provides peak values of the states [9], [10].

To the best of our knowledge, the Error in Variables (EIV) case has not been addressed in the context of data-driven control. Writing (1) in terms of the measured variables,

$$\hat{x}_{t+1} - \Delta x_{t+1} = A(\hat{x}_t - \Delta x_t) + B(\hat{u}_t - \Delta u_t) + w_t, \tag{3}$$

highlights the main difficulty here: the bilinearities $(A\Delta x_t,\ B\Delta u_t)$ between unknown variables that lead to generically NP hard problems. This paper proposes a convex, computationally tractable convex relaxation for robust data driven control with ℓ_∞ bounded measurement and process noise. Its main contributions are

1) To show that, in this scenario, robust superstabilizing controllers can be designed by solving Sum of Squares (SOS)-based feasibility problems, which can be posed as Semidefinite Programs (SDPs). Robust stabilization is guaranteed by ensuring that all closed loop plants consistent with the observed data are superstable.
2) A theorem of alternatives reformulation that drastically reduces the number of variables involved and yields more tractable SDPs [11].

This paper has the following structure: Section II will review preliminaries such as notation, notions of stability for linear systems, and SOS proofs of polynomial nonnegativity. Section III will present a description of the semialgebraic consistency sets and an SOS program in variables $(A, B, \Delta x)$ to find a superstabilizing controller. Section IV applies a Theorem of Alternatives to form an equivalent SOS program in (A, B) with a reduced computational complexity as compared to the Full program in $(A, B, \Delta x)$. Section V performs a comparison of computational complexity between the Full and Alternatives program. Section VI presents numerical experiments validating this method. Section VII details extensions such as the varying noise sets, input noise,

[1]J. Miller, T. Dai, and M. Sznaier are with the Robust Systems Lab, ECE Department, Northeastern University, Boston, MA 02115. (e-mails: {miller.jare, dai.ti}@northeastern.edu, msznaier@coe.neu.edu).

J. Miller, T. Dai, and M. Sznaier were partially supported by NSF grants CNS–1646121, ECCS–1808381 and CNS–2038493, AFOSR grant FA9550-19-1-0005, and ONR grant N00014-21-1-2431. 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, and the AFOSR International Student Exchange Program.

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9992363

and the combination of process noise and EIV. The paper is concluded in Section VIII.

II. Preliminaries

A. Notation

The set $\mathbb{R}^n$ is the n-dimensional Euclidean space, $\mathbb{R}^n_+$ is the nonnegative real orthant, and the set $\mathbb{R}^{m\times n}$ is the set of $m \times n$ matrices with real number entries. The set of real polynomials with indeterminate variables x is $\mathbb{R}[x]$, and the set of polynomials up to degree d is $\mathbb{R}[x]_{\leq d}$. The notation $(\mathbb{R}[x])^{m\times n}$ will correspond to a matrix-valued $m\times n$ polynomial with a variable x, and $(\mathbb{R}[x])^m$ is a vector-valued $m \times 1$ polynomial. The transpose of a matrix M is M^T, and a square symmetric matrix ($M = M^T$) is Positive Semidefinite (PSD) ($M \succeq 0$) if $x^T Mx \geq 0$ for all $x \neq 0$. The ℓ_∞ operator norm of a matrix M is $\|M\|_\infty = \max_j |M_{ij}|$. The imaginary number is $\mathbf{j} = \sqrt{-1}$, and the symbol $\mathbf{1}$ is a vector of all ones. The set of natural numbers between 1 and N is $1..N$.

B. Superstability of Discrete-Time Linear Systems

A closed loop system $A_{cl} = A + BK$ is superstable if,

$$\|A + BK\|_\infty < 1 \quad (\ell_\infty \text{ Operator Norm}). \tag{4}$$

Superstability implies that the ℓ_∞ norm $\|x\|_\infty$ is a polyhedral Lyapunov function of the closed loop system, proving that the origin is globally asymptotically stable. Another consequence of superstability is that every pole $p_i = a_i + \mathbf{j}b_i$ of A_{cl} satisfies $|a_i| + |b_i| < 1$ (ℓ_1 norm of poles).

An equivalent definition of superstability through the method of convex lifts from [12] is that $\exists M \in \mathbb{R}^{n\times n}$ with,

$$\textstyle\sum_{j=1}^n M_{ij} < 1 \qquad \forall i = 1..n \tag{5a}$$

$$-M_{ij} \leq A_{ij} + \textstyle\sum_{\ell=1}^m B_{i\ell}K_{\ell j} \leq M_{ij} \quad \forall i,j = 1..n. \tag{5b}$$

If $A + BK$ is superstable, an admissible selection of M satisfying (5) is $M_{ij} = |A_{ij} + \sum_{\ell=1}^m B_{i\ell}K_{\ell j}|,\ \forall i,j = 1..n$. Superstability is not necessarily preserved under a change-of-basis transformation of the closed-loop plant A_{cl}.

C. Semialgebraic Geometry and Sum of Squares

A Basic Semialgebraic (BSA) set is a set defined by a finite number of bounded-degree inequality and equality constraints. Every BSA set $\mathbb{K}$ can be represented as,

$$\mathbb{K} = \{x \mid g_i(x) \geq 0,\ h_j(x) = 0\}, \tag{6}$$

for appropriate describing polynomials $\{g_i(x)\}_{i=1}^{N_g}$ and $\{h_j(x)\}_{j=1}^{N_h}$. The intersection of two BSA sets remains BSA, and may be acquired by concatenating the describing polynomial constraints. The projection operation $\pi^x : (x,y) \mapsto x$ applied to a BSA set $\bar{\mathbb{G}}(x,y)$ is,

$$\mathbb{G}(x) = \pi^x \bar{\mathbb{G}}(x,y) = \{x \mid \exists y : (x,y) \in \bar{\mathbb{G}}\}. \tag{7}$$

Semialgebraic sets are the closure of BSA sets under unions and projections. The projections of BSA sets in (x,y) may be described as the union of disjoint BSA sets in x alone, and this task may be accomplished through quantifier elimination algorithms such as the Cylindrical Algebraic Decomposition [13] in typically (doubly) exponential time.

A polynomial nonnegativity constraint for $p(x) \in \mathbb{R}[x]$ is $p(x) \geq 0, \forall x \in \mathbb{K}$. Verifying polynomial nonnegativity is generically NP-hard, but SOS methods employ SDPs to find nonnegativity certificates through convex means [14]. A polynomial $p(x)$ is SOS ($p(x) \in \Sigma[x]$) if there exists a vector of polynomials $v(x) \in (\mathbb{R}[x])^s$ and a symmetric PSD matrix $Q \in \mathbb{R}^{s\times s}$ such that $p(x) = v(x)^T Qv(x)$. The Q matrix is also called the *Gram* matrix. If $Q = S^T S$ is a matrix decomposition of Q, then the elements $q(x) = Sv(x)$ satisfy $p(x) = \sum_{i=1}^s q_i(x)^2$. The vector v is often chosen as a monomial map up to a specified degree, where there exists $\binom{n+d}{d}$ monomials in n variables up to degree d.

The Putinar Positivestellensatz (Psatz) gives a condition for a polynomial $p(x)$ to be positive over a BSA $\mathbb{K}$ [15],

$$p(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x)g_i(x) + \sum_j \phi_j(x)h_j \tag{8a}$$

$$\exists\sigma_0(x) \in \Sigma[x], \quad \sigma(x) \in (\Sigma[x])^{N_g}, \quad \phi \in (\mathbb{R}[x])^{N_h}. \tag{8b}$$

The set $\mathbb{K}$ is *Archimedean* if there exists an $R \in (0,\infty)$ such that $R - \|x\|_2^2$ has a Putinar certificate in the sense of (8). The set of polynomials in (8b) is called the Weighted Sum of Squares (WSOS) cone $\Sigma[\mathbb{K}]$. The degree-($\leq$ d) WSOS cone $\Sigma[\mathbb{K}]_{\leq 2d}$ restricts all polynomials $(\sigma_0(x), \{\sigma_i(x)g_i(x)\}_{i=1}^{N_g}, \{\phi_j(x)h_j(x)\}_{j=1}^{N_h})$ to have degree at most $2d$. If the set $\mathbb{K}$ is Archimedean, then for every bounded-degree $p(x)$ that is positive over $\mathbb{K}$, there exists a finite integer d such that $p(x) \in \Sigma[\mathbb{K}]_{\leq 2d}$. The process of increasing the degree until a WSOS certificate is found is called the (moment)-SOS hierarchy, and each step in the hierarchy requires solving an SDP of increasing complexity. Details about the convergence rate of the moment-SOS hierarchy for polynomial optimization problems as d increases may be found in [16].

The per-iteration complexity of an Interior Point Method in solving (up to arbitrary accuracy) an SDP with M affine constraints and a PSD constraint of size N is $O(N^3M + M^2N^2)$ [17]. Finding a degree-d SOS certificate of $p(x)$'s positivity over $\mathbb{R}^n$ requires a Gram matrix Q of size $N = \binom{n+d}{d}$ and a set of $M = \binom{n+2d}{2d}$ affine constraints between coefficients of p and sums of coefficients in Q. The per-iteration complexity of Putinar-derived SOS SDPs therefore scales in a polynomial manner as d increases for fixed n as $O(d^{4n})$, and *vice versa* as n increases for fixed d as $O(n^{6d})$.

III. Superstabilizing Controller Design via SOS

This section will present an SOS feasibility program to recover a superstabilizing controller K compatible with all plants consistent with $\mathcal{D} \doteq \{\hat{x}_t, \hat{u}_t\}_{t=1}^T$ and the noise bounds. For simplicity, we start with the case where $w_t, \Delta u_t \equiv 0$ and defer the analysis where these input and process noise terms are present to Section VII.

A. Consistency Sets

The BSA set of plants $(A,B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$ and noise values $\Delta x \in \mathbb{R}^{n\times T}$ consistent with $\mathcal{D}$ and noise bound ϵ is the set $\bar{\mathcal{P}}(A,B,\Delta x)$ such that:

$$\bar{\mathcal{P}}: \begin{Bmatrix} 0 = -\Delta x_{t+1} + A\Delta x_t + h^0_t & \forall t = 1..T-1 \\ \|\Delta x_t\|_\infty \leq \epsilon & \forall t = 1..T \end{Bmatrix}, \quad (9)$$

where the affine weight h^0 is defined by,

$$h^0_t = \hat{x}_{t+1} - A\hat{x}_t - Bu_t \qquad \forall t = 1..T-1. \quad (10)$$

Remark 1: Data from multiple trajectories of the same system $\{\mathcal{D}_k\}_{k=1}^{N_d}$ may be merged to form $\bar{\mathcal{P}} = \cap_{k=1}^{N_d} \bar{\mathcal{P}}(\mathcal{D}_k)$.

The semialgebraic consistency set of plants $\mathcal{P}(A, B)$ compatible with $\mathcal{D}$ is the projection,

$$\mathcal{P}(A, B) = \pi^{A,B}\bar{\mathcal{P}}(A, B, \Delta x). \quad (11)$$

B. Statement of the Problem

Given an a-priori bound ϵ on the ℓ_∞ norm of the noise and experimental data $\mathcal{D}$, our goal is to find a gain K such that the closed loop system $(A + BK)$ is superstable for all pairs (A, B) in the consistency set. Formally:

Problem 1: Find K such that $\|A + BK\|_\infty < 1$, for all $(A, B) \in \mathcal{P}$.

Remark 2: All constraints describing $\bar{\mathcal{P}}$ in (9) are affine in the noise terms Δx. For a fixed plant (A_0, B_0), checking set membership $(A_0, B_0) \in \mathcal{P}$ can be determined by solving a Linear Program (LP) feasibility problem in Δx.

Remark 3: The sets $\bar{\mathcal{P}}$ and $\mathcal{P}$ may be disconnected.

C. An Equivalent Nonnegativity Program

Superstabilization of all plants in $\mathcal{P}$ by a given controller $K \in \mathbb{R}^{m\times n}$ can be certified through equation (5). The M matrix may be chosen as a matrix-valued function $M(A, B, \Delta x) : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times T} \to \mathbb{R}^{n\times n}$ that can vary over plants (A, B) and noise in the consistency set Δx and satisfies:

$$\forall i = 1..n : 1 - \delta - \textstyle\sum_{j=1}^n M_{ij}(A, B, \Delta x) \geq 0 \quad (12a)$$

$$\begin{aligned} &\forall i = 1..n,\ j = 1..n : \quad (12b) \\ &\quad M_{ij}(A, B, \Delta x) - (A_{ij} + \textstyle\sum_{\ell=1}^m B_{i\ell}K_{\ell j}) \geq 0 \\ &\quad M_{ij}(A, B, \Delta x) + (A_{ij} + \textstyle\sum_{\ell=1}^m B_{i\ell}K_{\ell j}) \geq 0 \end{aligned}$$

for some sufficiently small stability margin $\delta > 0$.

The following assumption is required for finite convergence of the sequence of relaxations to Problem 1,

Assumption 1 (Compactness): Sufficient data is collected such that $\bar{\mathcal{P}}$ (and therefore $\mathcal{P}$) are compact (Archimedean).

Lemma 3.1: The function $M(A, B, \Delta x)$ has a continuous selection under Assumption 1.

Proof: (sketch) For fixed K, let S be the set of feasible $M \in \mathbb{R}^{n\times n}$ satisfying (12). The set-valued map $\Xi_K : \bar{\mathcal{P}} \rightsquigarrow S$ is lower semicontinuous by Thm. 2.2 of [18]. The map Ξ_K has closed and convex images in S, so by Michael's Theorem (9.1.2 in [19]), a continuous selection exists. ∎

Lemma 3.2: The function $M(A, B, \Delta x)$ can be taken to be a polynomial $M_p(A, B, \Delta x)$.

Proof: (sketch) Assumption 1 and continuity of the $M(A, B, \Delta x)$ allows to find a polynomial approximation $M_p(A, B, \Delta x)$ satisfying (12) by invoking the Stone-Weierstrass theorem [20]. ∎

Using Lemma 3.2, Problem 1 can be recast into the following polynomial feasibility form:

Problem 2: Find K and a polynomial matrix $M(A, B, \Delta x)$ such that (12) holds for all $(A, B, \Delta x) \in \bar{\mathcal{P}}(A, B, \Delta x)$.

D. SOS Program and Numerical Considerations

Program (12) may be approximated through SOS methods as discussed in Section II-C by imposing that M is a polynomial matrix $M(A, B, \Delta x) \in (\mathbb{R}[A, B, \Delta x])^{n\times n}$.

Let $q^{\text{row}}_i(A, B, \Delta x; K)$ be the LHS constraint of equation (12a), and $q^{\pm}_{ij}(A, B, \Delta x; K)$ be the LHS constraints of (12b).

As an example, one of the constraints from (12b) at (i, j) may be represented as

$$q^+_{ij}(A, B, \Delta x; K) = M_{ij}(A, B, \Delta x) - (A_{ij} + \textstyle\sum_\ell B_{i\ell}K_{\ell j}).$$

The degree-d WSOS tightening of Problem 2 is presented in Algorithm 1 (up to > 0 in (8) and ≥ 0 in (12)).

Algorithm 1: Full Superstability Program

Input: $d,\ \delta,\ \mathcal{D},\ \epsilon$

Output: $K,\ M$ (or Infeasibility)

Solve (or find infeasibility certificate):

$$K \in \mathbb{R}^{n\times m} \quad (13a)$$

$$M \in (\mathbb{R}[A, B, \Delta x])^{n\times n}_{\leq 2d} \quad (13b)$$

$$q^{row}_i \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} \qquad \forall i \in 1..n \quad (13c)$$

$$q^{\pm}_{ij} \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} \qquad \forall i, j \in 1..n \quad (13d)$$

IV. Alternatives Program

While in principle Problem 2 can be solved using the techniques outlined above, the resulting SOS scales as $\binom{n(n+m+T)+d}{d}$, limiting the approach to relatively low order systems and short data records. This section addresses this issue by eliminating the noise variables Δx through the use of the Theorem of Alternatives.

A. Theorem of Alternatives

If the constraint,

$$q(A, B) \geq 0 \qquad \forall (A, B, \Delta x) \in \bar{\mathcal{P}}, \quad (14)$$

is satisfied, then the problem of finding an $(A, B, \Delta x) \in \bar{\mathcal{P}}$ with $-q(A, B) > 0$ is infeasible. Dual variable functions $\zeta^{\pm}(A, B) : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times T}_+$ and $\mu_{i,t}(A, B) : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times(T-1)}$ multiplying against the constraints in (9) may be defined for each fixed (A, B) to form the weighted sum,

$$\begin{aligned} S = &-q(A, B) + \textstyle\sum_{t=1}^T (\epsilon\mathbf{1} - \Delta x_t)^T\zeta^+_t + (\epsilon\mathbf{1} + \Delta x_t)^T\zeta^-_t \\ &+ \textstyle\sum_{t=1}^{T-1} \mu_t^T(-\Delta x_{t+1} + A\Delta x_t + h^0_t) \\ = &-q(A, B) + \textstyle\sum_{t=1}^T \epsilon\mathbf{1}^T(\zeta^+_t + \zeta^-_t) + \textstyle\sum_{t=1}^{T-1} \mu_t^T h^0_t \\ &+ \textstyle\sum_{t=1}^{T-1} \mu_t^T A\Delta x_t - \textstyle\sum_{t=2}^T \mu_t^T \Delta x_{t-1}. \end{aligned} \quad (15)$$

The terms of (15) that are independent of Δx may be isolated into $Q(A, B; \zeta^{\pm}, \mu)$ as,

$$Q = -q(A,B) + \textstyle\sum_{t=1}^{T} \epsilon \mathbf{1}^T(\zeta^+_{t,i} + \zeta^-_{t,i}) + \sum_{t=1}^{T-1} \mu_t^T h_t^0. \quad (16)$$

Finding a $(\zeta^{\pm}, \mu)$ pair such that $\sup_{\Delta x \in \mathbb{R}^{n \times T}} S \leq 0$ is necessary and sufficient to prove that (14) holds (by [11] and Section 5.8 of [21]), given that the describing constraints in (9) are affine (convex and concave) in Δx. The supremal value of S for each $(A, B; \zeta^{\pm}, \mu)$ is,

$$\sup_{\Delta x} S = \begin{cases} Q & \zeta_1^+ - \zeta_1^- = A^T \mu_1 \\ & \zeta_T^+ - \zeta_T^- = -\mu_{T-1} \\ & \zeta_t^+ - \zeta_t^- = A^T \mu_t - \mu_{t-1} \ \forall t = 2..T-1 \\ & \zeta^{\pm}_{1:T} \geq 0 \\ \infty & \text{else.} \end{cases} \quad (17)$$

The term $Q(A, B; \zeta^{\pm}, \mu)$ must be nonpositive, and the case statements on the right side of (17) must be valid in order for the supremal S to be nonpositive. An equivalent statement to the nonnegativity constraint in (14) is that,

$$\begin{aligned} &\exists\ \zeta^{\pm}_{1:T}(A,B) \geq 0,\ \mu_{1:T-1}(A,B): && (18a)\\ &Q(A, B; \zeta^{\pm}, \mu) \leq 0 \qquad \forall (A,B) \in \mathcal{P} && (18b)\\ &\zeta_1^+ - \zeta_1^- = A^T \mu_1 && (18c)\\ &\zeta_t^+ - \zeta_t^- = A^T \mu_t - \mu_{t-1} \qquad \forall t \in 2..T-1 && (18d)\\ &\zeta_T^+ - \zeta_T^- = -\mu_{T-1}. && (18e) \end{aligned}$$

Remark 4: The multipliers $(\zeta^{\pm}, \mu)$ have continuous selections in the compact $\mathcal{P}$ by similar arguments to Lemma 3.1.

B. Alternatives for Superstabilization

A new assumption is required to provide convergence guarantees in the SOS hierarchy associated with (18),

Assumption 2 (Archimedean): An Archimedean set $\Pi(A, B) \supseteq \mathcal{P}$ is a-priori known.

Remark 5: A set Π may arise from prior knowledge about plant behavior and its reasonable limits.

The WSOS formulation of the Alternatives certificate (18) for a single constraint (14) at degree d is an SDP with decision variables $(\zeta^{\pm}, \mu)$ as described in Algorithm 2. The notation $\Sigma^{\text{altern}}_{\leq 2d}[\mathcal{P}]$ will refer to the cone of functions $q(A, B)$ with certificates given by (19) at degree d.

Remark 6: Constraints (19d)-(19e) are a set of linear inequality constraints in the coefficients of $(\zeta^{\pm}, \mu)$ with respect to indeterminates (A, B).

Remark 7: The multipliers $\zeta^{\pm}$ may be degree $2d$, since they are no longer Psatz multipliers in (8) against constraints $\epsilon \pm \Delta x_{it}$. The multipliers μ have degree $2d-1$ to ensure that the product $A^T \mu_t$ in (19d)-(19e) has degree $2d$.

Algorithm 3 for Alternatives-based superstabilization replaces each of the $2n^2 + n$ Putinar Psatz (8) calls in (13c)-(13d) with the Alternatives Psatz (19) in (20c)-(20d).

Remark 8: Assumption 2 is necessary to assure convergence of certificate (20) as the degree d increases to the finite recovery value (with M independent of Δx). Dropping Assumption 2 may lead to valid superstabilizing K with certificates, but such programs do not possess a convergence guarantee as d increases.

Algorithm 2: Alternatives Psatz $(\Sigma^{\text{altern}}_{\leq 2d}[A, B])$

Input: $d,\ q(A,B), \Pi, \mathcal{D}, \epsilon$
Output: $\zeta,\ \mu$ (or Infeasibility)
Solve (or find infeasibility certificate):

$$\begin{aligned} &\zeta^{\pm}(A,B) \in (\Sigma[\Pi]_{\leq 2d})^{n \times T} && (19a)\\ &\mu(A,B) \in (\mathbb{R}[A,B]_{\leq 2d-1})^{n \times (T-1)} && (19b)\\ &-Q(A, B; \zeta^{\pm}, \mu) \in \Sigma[\Pi]_{\leq 2d} \text{ (from (16))} && (19c)\\ &\zeta_1^+ - \zeta_1^- = A^T \mu_1 && (19d)\\ &\zeta_t^+ - \zeta_t^- = A^T \mu_t - \mu_{t-1} \qquad \forall t \in 2..T-1 && (19e)\\ &\zeta_T^+ - \zeta_T^- = -\mu_{T-1}. && (19f) \end{aligned}$$

Algorithm 3: Alternatives Superstability Program

Input: $d,\ \delta,\ \mathcal{D},\ \epsilon,\ \Pi$
Output: $K,\ M$ (or Infeasibility)
Solve (or find infeasibility certificate):

$$\begin{aligned} &K \in \mathbb{R}^{n \times m} && (20a)\\ &M \in (\mathbb{R}[A,B])^{n \times n}_{\leq 2d} && (20b)\\ &q_i^{row} \in \Sigma^{\text{altern}}_{\leq 2d}[A,B] \qquad \forall i \in 1..n && (20c)\\ &q_{ij}^{\pm} \in \Sigma^{\text{altern}}_{\leq 2d}[A,B] \qquad \forall i,j \in 1..n && (20d) \end{aligned}$$

V. Computational Complexity

This section will quantify the decrease in computational complexity obtained when using the Alternatives program as compared to the Full method. From (12), we have $2n^2 + n$ scalar polynomials $q(A, B, \Delta x) \in \Sigma[x]_{2d}$ in $p_F = n(n+m+T)$ variables $(A, B, \Delta x)$, where the size of each polynomials is computed from (8). The notation $s(\cdot)$ stands for the size of vector $\mathbb{R}^{s(\cdot) \times 1}$ and $m(\cdot)$ stands for the size of matrix $\mathbb{R}^{m(\cdot) \times m(\cdot)}$:

Full	q	σ_0	σ_i	μ_j
# polys.	1	1	$2nT$	$2n(T-1)$
size	$s\binom{p_F+2d}{2d}$	$m\binom{p_F+d}{d}$	$m\binom{p_F+d-1}{d-1}$	$s\binom{p_F+2d-2}{2d-2}$

TABLE I: Size of Full method

Similarly from (14) and (8) we get the size of Alternatives method with $p_A = n(n+m)$ variables (A, B). :

Alternatives	q	σ_0	σ_i	μ_j
# polys.	1	1	$2nT$	$2n(T-1)$
size	$s\binom{p_A+2d}{2d}$	$m\binom{p_A+d}{d}$	$m\binom{p_A+d}{d}$	$s\binom{p_A+2d-1}{2d-1}$

TABLE II: Size of Alternatives method

Two major sources of complexity reduction are:
(a) in the Alternatives method, the number of variables p

does not depend on the number of samples T.
(b) We would like to use the smallest d such that the algorithm is feasible. Experimental results shows that Full method only works with $d \geq 2$ while the Alternatives method works with $d \geq 1$.

Remark 9: The multipliers μ against consistency constraints (3) have degree $2d-2$ in Full (bilinearity $A\Delta x$) but have degree $2d-1$ in Alternatives (affine in (A, B) after eliminating Δx).
TABLE III shows the size (not multiplicites) of the variables with fixed $n = 2, m = 1, d_{full} = 2, d_{altern} = 1$ and increased T.

	q	σ_0	σ_i	μ_j
Alternatives	28	7	7	7
Full (T = 4)	3060	120	15	120
Full (T = 6)	7315	190	19	190
Full (T = 8)	14950	276	23	276

TABLE III: Size of variables

VI. Numerical Examples

MATLAB (2021a) code to generate the examples below is publicly available at https://github.com/jarmill/error_in_variables. Dependencies include Mosek [22] and YALMIP [23].

A. Model-Based and Data-Driven Comparison

The model-based approach, (i.e. with A, B known) is formulated as the following program:

$$\min_{\lambda \in [0,1),\ K} \quad \lambda : \quad ||A + BK||_\infty < \lambda. \tag{21}$$

Here λ is a scalar variable representing the convergence rate. By minimizing λ, we obtain the fastest closed-loop system. Consider the following unstable discrete-time model:

$$A = \begin{bmatrix} 0.6852 & 0.0274 & 0.5587 \\ 0.2045 & 0.6705 & 0.1404 \\ 0.8781 & 0.4173 & 0.1981 \end{bmatrix}, B = \begin{bmatrix} 0.4170 & 0.3023 \\ 0.7203 & 0.1468 \\ 0.0001 & 0.0923 \end{bmatrix} \tag{22}$$

We excite the system with uniformly distributed input and measurement noise with bound $||u||_\infty = 1$, $||\Delta x_t||_\infty = \epsilon$. The initial state is $x_1 = [1, 0, 0]$. A trajectory of T samples, i.e. $\{\hat{x}_t, u_t\}_{t=1}^T$ is collected for design. Solving (21) with known A, B from (22) leads to $\lambda_{true} = 0.7259$. We treat this as a benchmark and compare this with λ obtained with the data-driven approach. To avoid the computational burden, we choose the lowest order for all examples, i.e. $d_{full} = 2$ and $d_{altern} = 1$. We drop Assumption 2 as noted in Remark 8 for the Alternatives program.
For a horizon $T = 6$ and $\epsilon = 0$ (clean data), the Full method introduces approximately 3.4×10^7 variables which is beyond the current capabilities of Mosek. On the other hand, the Alternatives method only has 67776 scalar variables (3 orders of magnitude smaller than the Full). Solving it leads to $\lambda = 0.7259 = \lambda_{true}$ which indicates that there is no conservatism in our data-driven method for clean data. Now we consider the noisy case with $\epsilon = 0.05$. Applying the algorithm with $T = 40$ leads to $\lambda = 0.8880$. Note that this λ corresponds to the worse-case convergence rate, i.e. the largest convergence rate for all plants in the consistency set. The true closed-loop convergence rate is obtained by computing the norm of the closed-loop system, which is, $\lambda_{clp} = ||A+BK||_\infty = 0.7749$. It is worth noting that $\lambda_{true} \leq \lambda_{clp} \leq \lambda$.

B. Monte Carlo Simulations

To test the reliability of the proposed method, we collected 100 trajectories with different level of noise and applied the Alternatives method to the following system:

$$A = \begin{bmatrix} 0.6863 & 0.3968 \\ 0.3456 & 1.0388 \end{bmatrix}, \quad B = \begin{bmatrix} 0.4170 & 0.0001 \\ 0.7203 & 0.3023 \end{bmatrix} \tag{23}$$

TABLE IV displays the number of successful designs (S) for a fixed horizon of $T = 8$.

ϵ	0.05	0.08	0.11	0.14
S	100	84	57	39

TABLE IV: S as a function of ϵ with $T = 8$

Increasing the noise level expands the consistency set, which in turn renders the problem of finding a single superstabilizing controller more difficult. Collecting more data with the same noise bound $\epsilon = 0.14$ reduces the size of the consistency set, as illustrated in TABLE V.

T	8	10	12	14
S	39	60	75	86

TABLE V: S as a function of T with $\epsilon = 0.14$

C. Partial Information

It is easy to incorporate partial information in the proposed framework. Instead of treating all entries of A, B as unknown variables, we can assume that 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 $\binom{n(n+m)+d}{d}$ and ensuring that it is easier to find a superstabilizing K both theoretically and computationally. For instance, if we assume that the second column of A is known and apply the alternative method with $T = 8$, $\epsilon = 0.14$, we get $S = 94$ as compared to $S = 39$ in the last column of TABLE IV.

VII. Extensions

This section sketches out various extensions to the presented nonnegativity-based superstabilization framework.

A. Varying Noise Sets

The constraint description for $\bar{\mathcal{P}}$ in equation (9) involves a noise bound of $\|\Delta x_t\| \leq \epsilon$ for each $t = 1..T$. Time-dependent noise constraints may be developed by defining sets $\mathcal{F}_t$ such that $\Delta x_t \in \mathcal{F}_t$. Algorithm 1 for Full stabilization will function when each $\mathcal{F}_t$ is BSA in Δx. The Alternatives psatz in Alg. 2 and its program in Alg. 3 may be adapted when $\mathcal{F}_t$ are polytopes.

B. Input Noise

The data $\mathcal{D}$ in this paper assumed bounded measurement noise in the state x (Δx) and perfect knowledge of the input u. Let $\|\Delta x_t\|_\infty \leq \epsilon_x$ and $\|\Delta u_t\|_\infty \leq \epsilon_u$ be measurement noise processes for the state and the input. Data $\mathcal{D} = (\hat{x}_t, \hat{u}_t)$ is now collected under the relation,

$$\hat{x}_t = x_t + \Delta x_t \qquad \hat{u}_t = u_t + \Delta u_t. \tag{24}$$

Relation (3) with added input noise is,

$$\hat{x}_{t+1} - \Delta x_{t+1} = A(\hat{x}_t - \Delta x_t) + B(\hat{u}_t - \Delta u_t). \tag{25}$$

The consistency sets $\bar{\mathcal{P}}$ and $\mathcal{P}$ may be defined with respect to the zero locus of relation (25). The full program with input noise will have $n(n+m) + T(n+m)$ variables. The Alternatives program (14) will involve multipliers $\zeta^\pm_x$ over Δx and $\psi^\pm_u$ over Δu. The term Q with input noise is $Q(A, B; \zeta^\pm, \psi^\pm, \mu) = Q(A, B; \zeta^\pm, \mu) + \sum_{t=1}^T \epsilon \mathbf{1}^T(\psi^+_t + \psi^-_t)$, and the new constraints when eliminating Δu are $\psi^+_t - \psi^-_t = B^T \mu_t, \ \forall t = 1..T$.

C. Process and Measurement Noise

Assume that the measurement and process noise have ℓ_∞ norm bounds of $\|\Delta x_t\|_\infty \leq \epsilon_x$ and $\|w_t\|_\infty \leq \epsilon_p$. The consistency set of plants and measurement noise $\bar{\mathcal{P}}^{\epsilon_x, \epsilon_w}$ is,

$$\begin{aligned} &(A, B, \Delta x) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n \times m} \times \mathbb{R}^{n \times T} : && (26)\\ &\quad \|-\Delta x_{t+1} + A\Delta x_t + h^0_t\|_\infty \leq \epsilon_p && \forall t = 1..T-1 \\ &\quad \|\Delta x_t\|_\infty \leq \epsilon && \forall t = 1..T \end{aligned}$$

Each of the $2n^2 + n$ nonnegativity expressions in (12) may be posed over the BSA set $\bar{\mathcal{P}}^{\epsilon_x, \epsilon_w}$ in (26) by the Putinar Psatz in (8). The full program with process and measurement noise still has $n(n + m + T)$, but there are $2n(T-1)$ additional inequality constraints arising from the process noise. The Alternatives method would no longer have μ multipliers against consistency equality constraints, instead the ϵ_p inequality constraints would have multipliers $\xi^\pm$. Each instance of μ in constraints (19d)-(19f) is replaced by $\xi^+ - \xi^-$, and the multiplier term $Q + q(A, B)$ is now,

$$\epsilon \mathbf{1}^T \left(\textstyle\sum_{t=1}^T (\zeta^+_t + \zeta^-_t) + (\sum_{t=1}^{T-1} \xi^+_t + \xi^-_t) \right). \tag{27}$$

VIII. Conclusion

This work presented a convergent WSOS program (Full) to perform superstabilization of EIV models. To the best of our knowledge, this is the first paper to address the EIV stabilization scenario. A theorem of alternatives was then applied to produce an equivalent problem (Alternatives) with substantially reduced complexity. This was accomplished by using duality to eliminate the noise variables. Efficacy of this method was demonstrated on example systems.

Future work includes expanding the set of stable controllers beyond superstability, producing worst-case-LQR optimal controllers K with respect to all consistent plants in $\mathcal{P}$, and designing output feedback controllers in the measurement noise setting.

References

[1] 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, 2019.

[2] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," Systems & Control Letters, vol. 54, no. 4, pp. 325–329, 2005.

[3] 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.

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

[5] J. Berberich, C. W. Scherer, and F. Allgöwer, "Combining Prior Knowledge and Data for Robust Controller Design," arXiv preprint arXiv:2009.05253, 2020.

[6] 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.

[7] T. Dai and M. Sznaier, "Data driven robust superstable control of switched systems," IFAC-PapersOnLine, vol. 51, no. 25, pp. 402–408, 2018.

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

[9] 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.

[10] 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.

[11] 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.

[12] M. Yannakakis, "Expressing combinatorial optimization problems by Linear Programs," Journal of Computer and System Sciences, vol. 43, no. 3, pp. 441–466, 1991.

[13] B. F. Caviness and J. R. Johnson, Quantifier Elimination and Cylindrical Algebraic Decomposition. Springer Science & Business Media, 2012.

[14] P. A. Parrilo, Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. California Institute of Technology, 2000.

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

[16] J. Nie and M. Schweighofer, "On the complexity of Putinar's Positivstellensatz," Journal of Complexity, vol. 23, no. 1, pp. 135–150, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0885064X0600080X

[17] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," SIAM J OPTIMIZ, vol. 5, no. 1, pp. 13–51, 1995.

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

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

[20] M. H. Stone, "Applications of the theory of boolean rings to general topology," Transactions of the American Mathematical Society, vol. 41, no. 3, pp. 375–481, 1937.

[21] S. Boyd, S. P. Boyd, and L. Vandenberghe, Convex Optimization. Cambridge university press, 2004.

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

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