

Superstabilizing Control of Discrete-Time ARX Models under Error in Variables

Jared Miller¹, Tianyu Dai¹, Mario Sznaier¹ *[†]

July 11, 2023

Abstract

This paper applies a polynomial optimization-based framework towards the superstabilizing control of an Autoregressive with Exogenous Input (ARX) model given noisy data observations. The recorded input and output values are corrupted with L-infinity-bounded noise where the bounds are known. This is an instance of Error in Variables (EIV) in which true internal state of the ARX system remains unknown. The consistency set of ARX models compatible with noisy data has a bilinearity between unknown plant parameters and unknown noise terms. The requirement for a dynamic compensator to superstabilize all consistent plants is expressed using polynomial nonnegativity constraints, and solved using sum-of-squares (SOS) methods in a converging hierarchy of semidefinite programs in increasing size. The computational complexity of this method may be reduced by applying a Theorem of Alternatives to eliminate the noise terms. The effectiveness of this method is demonstrated on control of example ARX models.

1 Introduction

Data Driven Control (DDC) is a class of algorithms that perform control of plants consistent with measured data without requiring a system identification step first (Hou and Wang, 2013). Recent design methods for state space models are based on Willem’s Fundamental Lemma (Berberich et al., 2020; De Persis and Tesi, 2020), Matrix S-Lemma (van Waarde et al., 2020), and/or a theorem of alternatives (Dai and Sznaier, 2018).

Most DDC applications involve structured process noise in dynamics. The Error in Variables (EIV) case is considerably less studied for control. Prior work about system identification and state estimation of EIV-affected systems includes (Cerone, 1993; Cerone et al., 2011; Norton, 1987; Söderström, 2018). Polynomial optimization for system identification of EIV Autoregressive with Exogenous Inputs (ARX) models has previously been considered in Chapter 6 of (Cheng, 2016) using rank-minimization, but the EIV control task was not previously addressed. This paper is a continuation of research started by the authors in (Miller et al., 2022a,b) for EIV control with full-state feedback.

Superstability is a computationally tractable restriction of stability used in control and static output feedback (Polyak and Halpern, 2001; Polyak and Shcherbakov, 2002). Superstability allows for the design of controllers with guaranteed performance bounds on the growth of the L_∞ -norm of the state/output (Sznaier and Bianchini, 1995), which was extended to the matrix transfer function case in (Chen et al., 2005). Data-driven superstabilization of ARX models under process noise was performed in (Cheng et al., 2015).

To the best of our knowledge, this paper is the first work that addresses output-feedback DDC with noisy input-output measurements.

The contributions of this work are,

*¹J. 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 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.

- Formulation and solution of EIV ARX superstabilization as a Polynomial Optimization Problem (POP) using Sum of Squares (SOS) methods.
- Application of a Theorem of Alternatives to form a POP where noise terms $(\Delta u, \Delta y)$ are eliminated.
- Accounting of the computational complexity to solve the Semidefinite Programs (SDPs) with and without utilizing the Theorem of Alternatives.

This paper has the following structure: Section 2 introduces preliminaries such as notation, ARX models, Superstability applied to ARX models, and SOS methods. Section 3 presents a POP for superstabilizing control involving the variables $(a, b, \Delta u, \Delta y)$. Section 4 utilizes a Theorem of Alternatives to create an equivalent POP in terms of (a, b) . Section 5 tabulates the computational complexity of these SDPs. Section 6 applies the derived methods towards example superstabilization problems. Section 7 introduces extensions such as process noise and Switched ARX models. Section 8 concludes the paper.

2 Preliminaries

ARX Autoregressive with Exogenous Inputs

BSA Basic Semialgebraic

DDC Data Driven Control

EIV Error in Variables

LP Linear Program

POP Polynomial Optimization Problem

PSD Positive Semidefinite

SDP Semidefinite Program

SOS Sum of Squares

WSOS Weighted Sum of Squares

2.1 Notation

The set of real numbers is \mathbb{R} and the n -dimensional real Euclidean vector space is \mathbb{R}^n . The set of integers between r and s is $r..s$. The n -dimensional nonnegative real orthant is \mathbb{R}_+^n . The set of $n \times m$ real-valued matrices is $\mathbb{R}^{n \times m}$. The transpose of a matrix M is M^T . The identity matrix is I , the all-ones matrix is $\mathbf{1}$, and the all-zeros matrix is $\mathbf{0}$. The L_1 -norm of a vector x is $\|x\|_1 = \sum_{i=1}^n |x_i|$. The L_∞ norm of a vector x is $\|x\|_\infty = \max_{i \in 1..n} |x_i|$. The vertical concatenation of vectors x and y is $[x; y]$. The lag operator λ applied to a sequence $\{x_t\}$ is $\lambda x_t = x_{t-1}$. The cross-correlation $x \star y$ between $\{x_t\}_{t=l}^k$ and $\{y_t\}_{t=m}^n$ is

$$(x \star y)_j = \sum_{i=m}^n x_{i+j-n+m-1} y_i, \quad j = l..k + n - m. \quad (1)$$

The cardinality of a finite set \mathcal{A} is $|\mathcal{A}|$. The set of natural numbers is \mathbb{N} , and the set of n -dimensional multi-indices is \mathbb{N}^n . Every polynomial $p(x)$ may be defined with respect to some set $\mathcal{A} \subset \mathbb{N}^n$ as $p(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$ with all coefficients $c_\alpha \neq 0$. The degree of a polynomial $p(x)$ is $\deg p$. The ring of real-valued polynomials with indeterminates x is $\mathbb{R}[x]$, and its subset of polynomials with degree d or less is $\mathbb{R}[x]_{\leq d}$. The notation $(\mathbb{R}[x])^k$ will denote a k -vector of polynomials in x . The coefficient vector of a polynomial $p(x)$ is $c = \mathbf{coeff}[p(x)]$.

2.2 ARX Models

An ARX model with input-output sequence (u, y) and parameters (a, b) such that $n_a > n_b$ obeys dynamics for times $t \in 1..T$:

$$y_t = -\sum_{i=1}^{n_a} a_i y_{t-i} + \sum_{i=1}^{n_b} b_i u_{t-i}. \quad (2)$$

The ARX model in (2) may be represented as a rational transfer function in the lag operator λ with

$$G(\lambda) = \frac{\sum_{i=1}^{n_b} b_i \lambda^i}{1 + \sum_{i=1}^{n_a} a_i \lambda^i} = \frac{B(\lambda)}{1 + A(\lambda)}. \quad (3)$$

The numerator B and denominator A of (3) are each polynomials in λ . The function (3) satisfies $B(0) = 0$ and is assumed to be strictly proper ($G(0) = 0$).

Let $C(\lambda)$ be a dynamic compensator with parameters (\tilde{a}, \tilde{b}) of length $(\tilde{n}_a, \tilde{n}_b)$ following the structure of (3), yielding polynomials $(\tilde{A}(\lambda), \tilde{B}(\lambda))$ with $C(\lambda) = \tilde{B}(\lambda)/(1 + \tilde{A}(\lambda))$. Application of $C(\lambda)$ in feedback with $G(\lambda)$ yields the following closed-loop transfer function $G_{cl}(\lambda)$:

$$\frac{G(\lambda)}{1 + G(\lambda)C(\lambda)} = \frac{B(\lambda)(1 + \tilde{A}(\lambda))}{(1 + A(\lambda))(1 + \tilde{A}(\lambda)) + B(\lambda)\tilde{B}(\lambda)}. \quad (4)$$

The closed-loop system parameter a^{cl} related to (4) is

$$a^{cl} = \mathbf{coeff}[(1 + A(\lambda))(1 + \tilde{A}(\lambda)) + B(\lambda)\tilde{B}(\lambda) - 1]. \quad (5)$$

A single ARX system in (2) is stable if the roots of $1 + A(\lambda)$ are contained within the exterior of the unit disc $\{\lambda \mid \sqrt{\operatorname{Re}(\lambda)^2 + \operatorname{Im}(\lambda)^2} > 1\}$. This stability may be verified through numerical computation of roots or with the Jury stability criterion (Ogata, 1995). These methods are computationally expensive to employ when designing stabilizing compensators C (as in (4)) such that (5) is stable, given that the methods all involve polynomial constraints on the entries of (\tilde{a}, \tilde{b}) .

2.3 Superstability

Superstability (Blanchini and Sznaier, 1995; Polyak and Halpern, 2001; Polyak and Shcherbakov, 2002) is a conservative notion of stability that possesses simpler computational properties. An ARX system is superstable if $\|a\|_1 < 1$. If the system satisfies $\|a\|_1 < \gamma$ for some value $\gamma \in [0, 1)$, then a bound on $|y_t|$ based on the initial condition given $u = 0$ may be found as $\forall t \geq 0 : |y_t| \leq \gamma^{t/n_a+1} \max_{t' \in 0..n_a-1} |x_{-t'}|$ (Theorem A.1.(a) of (Polyak and Shcherbakov, 2002) when starting at time $t = 1$). The superstability norm constraint $\|a^{cl}\|_1 < 1$ from (5) is affine-expressible in the compensator parameters (\tilde{a}, \tilde{b}) .

Superstability may be imposed for classes of plants. A single controller C can simultaneously superstabilize a set (a, b) of plants if the closed-loop system (4) is superstable for each individual plant.

2.4 Sum of Squares

A Basic Semialgebraic (BSA) set \mathbb{K} is the locus of a finite number of bounded-degree polynomial inequality and equality constraints:

$$\mathbb{K} = \{x \in \mathbb{R}^n \mid g_k(x) \geq 0, h_{k'}(x) = 0\}, \quad (6)$$

for all indices $k = 1..N_i$, $k' = 1..N_e$. BSA sets may be intersected by concatenating their describing polynomials $\{g\}$ and $\{h\}$. The x -projection operator $\pi^x : X \times Y \rightarrow X$ is $\pi^x : (x, y) \mapsto x$. A BSA set $\mathbb{K} \subset X \times Y$ has an x -projection $\pi^x \mathbb{K} = \{x \in X \mid (x, y) \in \mathbb{K}\}$. BSA sets are not closed under projection, instead the projection of a BSA set is generically the union of disjoint BSA sets.

A polynomial is nonnegative if $\forall x \in \mathbb{R}^n : p(x) \geq 0$. A polynomial $p(x)$ is SOS if there exists an integer s , a polynomial vector $v \in (\mathbb{R}[x])^s$, and an $s \times s$ Positive Semidefinite (PSD) Gram matrix $Q \succeq 0$ such that $p(x) = v(x)^T Q v(x)$. SOS polynomials are a subset of nonnegative polynomials, given that the square

of any real number is nonnegative. The set of SOS polynomials is $\Sigma[x] \subset \mathbb{R}[x]$, and its subset of degree- $2d$ polynomials is $\Sigma[x]_{\leq 2d} \subset \mathbb{R}[x]_{\leq 2d}$ (SOS polynomials are always even).

A sufficient condition for a polynomial $p(x)$ to be positive over a BSA set \mathbb{K} is (Putinar, 1993)

$$p(x) = \sigma_0(x) + \sum_i \sigma_i(x)g_i(x) + \sum_j \phi_j(x)h_j \quad (7a)$$

$$\exists \sigma_0(x) \in \Sigma[x], \quad \sigma(x) \in (\Sigma[x])^{N_g}, \quad \phi \in (\mathbb{R}[x])^{N_h}. \quad (7b)$$

The Weighted Sum of Squares (WSOS) set $\Sigma[\mathbb{K}]$ is the set of polynomials that admit a certificate in (7) (called a Putinar Psatz). The Putinar certificate in (7) is necessary and sufficient if an *Archimedean* condition holds: $\exists R > 0 \mid R - \|x\|_2^2 \in \Sigma[\mathbb{K}]$. Every Archimedean BSA set is compact, and compact sets may be rendered Archimedean by adjoining $R - \|x\|_2^2 \geq 0$ to the descriptor constraints if a valid R is previously known. The moment-SOS hierarchy for optimization involves increasing the degree $2d$ to obtain higher-order Putinar multipliers (7b) (Lasserre, 2009). Certifying that a degree- $2d$ polynomial $p(x)$ is SOS requires $1 + N_g$ Gram matrices of maximal size $\binom{n+d}{d}$, and the approximate per-iteration runtime of an SDP originating from the moment-SOS hierarchy is approximately $O(n^{6d})$ when d is fixed.

3 Superstabilizing Control

This section presents a POP to perform superstabilizing control of EIV ARX models.

3.1 Problem Statement

A set of input-output observations $\mathcal{D} = (\hat{u}, \hat{y})$ are recorded for an ARX dynamical system within a time horizon of T . The ground-truth is the input-output data (u, y) following dynamics (2). The records in \mathcal{D} are corrupted by L_∞ -bounded input noise ($\Delta u \in \mathbb{R}^{T+n_b-1}$) and measurement noise ($\Delta y \in \mathbb{R}^{T+n_a}$) as

$$\hat{y} = y + \Delta y, \quad \|\Delta y\|_\infty \leq \epsilon_y \quad (8a)$$

$$\hat{u} = u + \Delta u, \quad \|\Delta u\|_\infty \leq \epsilon_u. \quad (8b)$$

The combination of input and measurement noise is the EIV setting. Substitution of (8) into (2) yields

$$\begin{aligned} \hat{y}_t - \Delta y_t &= \left(- \sum_{i=1}^{n_a} a_i \hat{y}_{t-i} + \sum_{i=1}^{n_b} b_i \hat{u}_{t-i} \right) \\ &\quad - \left(- \sum_{i=1}^{n_a} a_i \Delta y_{t-i} + \sum_{i=1}^{n_b} b_i \Delta u_{t-i} \right). \end{aligned} \quad (9)$$

3.2 Consistency Sets

The $(\Delta u, \Delta y)$ -constant terms in (9) may be written as

$$h_t = \hat{y}_t + \sum_{i=1}^{n_a} a_i \hat{y}_{t-i} - \sum_{i=1}^{n_b} b_i \hat{u}_{t-i}. \quad (10)$$

Eq. (9) may therefore be expressed in terms of h_t as

$$0 = h_t + \left(- \sum_{i=1}^{n_a} a_i \Delta y_{t-i} + \sum_{i=1}^{n_b} b_i \Delta u_{t-i} \right) - \Delta y_t. \quad (11)$$

The set of parameters (a, b) and noise values $(\Delta u, \Delta y)$ consistent with data \mathcal{D} is

$$\bar{\mathcal{P}} : \left\{ \begin{array}{l} a \in \mathbb{R}^{n_a}, b \in \mathbb{R}^{n_b} \\ \Delta u \in \mathbb{R}^{T+n_b-1} \\ \Delta y \in \mathbb{R}^{T+n_a} \end{array} \left| \begin{array}{l} \|\Delta u\|_\infty \leq \epsilon_u \\ \|\Delta y\|_\infty \leq \epsilon_y \\ \text{Eq. (11)} \forall t = 1..T \end{array} \right. \right\}. \quad (12)$$

The set of parameters (a, b) consistent with data in \mathcal{D} is

$$\mathcal{P}(a, b) = \pi^{a,b} \bar{\mathcal{P}}(a, b, \Delta u, \Delta y). \quad (13)$$

Equivalently, a plant (a, b) is a member of \mathcal{P} if there exists an admissible L_∞ -bounded noise process $(\Delta u, \Delta y)$ that could have generated \mathcal{D} .

The following assumption will be required to obtain convergence:

A1 The set $\bar{\mathcal{P}}$ is compact.

Remark 1. Compactness of $\bar{\mathcal{P}}$ by A1 implies that its projection \mathcal{P} is also compact.

Remark 2. Determining membership (if a fixed plant $(a, b) \in \mathcal{P}$) is a Linear Program (LP) with variables $\Delta u, \Delta y$, given that $\bar{\mathcal{P}}$ has a bilinear description.

3.3 Full Program

The coefficients a^{cl} have maximal length $n_{cl} = \tilde{n}_a + n_a$.

Problem 1. A program to perform superstabilizing control (or to find an infeasibility certificate if $\gamma^* \geq 1$) is

$$\gamma^* = \min_{\gamma \in \mathbb{R}, (\tilde{a}, \tilde{b})} \gamma \quad \forall (a, b, \Delta u, \Delta y) \in \bar{\mathcal{P}} : \gamma \geq \|a^{cl}\|_1. \quad (14a)$$

The L_1 -norm constraint in (14) may be equivalently represented by a lifted sequence of inequalities. The L_1 -norm of a vector $x \in \mathbb{R}^n$ may also be expressed as $\|x\|_1 = \min_{m \in \mathbb{R}^{n_{cl}}} \sum_i m_i : -m_i \leq x_i \leq m_i$ (Gouveia et al., 2013). Letting $m_i(a, b) : \mathcal{P} \rightarrow \mathbb{R}$ be a set of functions for $i = 1..n_a + \tilde{n}_a$, expression (14) may be written as

$$\gamma^* = \min_{\gamma, (\tilde{a}, \tilde{b})} \gamma \quad (15a)$$

$$\forall (a, b, \Delta u, \Delta y) \in \bar{\mathcal{P}} : \quad (15b)$$

$$\gamma - \sum_{i=1}^{n_{cl}} m_i(a, b) \geq 0 \quad (15c)$$

$$m_i(a, b) - a_i^{cl}(a, b) \geq 0 \quad \forall i = 1..n_{cl} \quad (15d)$$

$$m_i(a, b) + a_i^{cl}(a, b) \geq 0 \quad \forall i = 1..n_{cl}. \quad (15e)$$

3.4 Full Sum-of-Squares Program

Problem 1 may be solved using SOS programming. The nonnegativity constraints in (15c)-(15e) may each be realized as Psatz constraints in the sense of (7). The degree- d polynomial restriction to Problem 1 is presented in Eq. (16). The decision variables of Eq. (16) are the L_1 -certificates m and the gain γ . The \geq symbols in (15) are tightened to $>$ in Eq. (16) due to the Putinar Psatz's (7) positivity certificate.

$$\gamma_d^* = \min_{\gamma \in \mathbb{R}} \gamma \quad (16a)$$

$$\tilde{a} \in \mathbb{R}^{n_a}, \tilde{b} \in \mathbb{R}^{n_b} \quad (16b)$$

$$m_i \in (\mathbb{R}[a, b])_{\leq 2d}^{n_{cl}} \quad \forall i \in 1..n_{cl} \quad (16c)$$

$$\gamma - \sum_{i=1}^{n_{cl}} m_i(a, b) \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} \quad (16d)$$

$$m_i(a, b) - a_i^{cl}(a, b) \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} \quad \forall i = 1..n_{cl} \quad (16e)$$

$$m_i(a, b) + a_i^{cl}(a, b) \in \Sigma[\bar{\mathcal{P}}]_{\leq 2d} \quad \forall i = 1..n_{cl}. \quad (16f)$$

Theorem 3.1. The functions m_i have continuous selections.

Proof. Let $\mathcal{M}_{\gamma, C} : \mathcal{P} \rightrightarrows \mathbb{R}^{n_{cl}}$ be the set-valued map defining the closed convex solution region $(a, b) \mapsto \{m \in \mathbb{R}^{n_{cl}} \mid \sum_{i=1}^{n_{cl}} m_i \leq \gamma, \forall i : m_i \geq \pm a_i^{cl}(a, b)\}$ from (15) given $(\gamma, \tilde{a}, \tilde{b})$. Note that the functions $a_i^{cl}(a, b)$ from (5) are linear (continuous) functions of (a, b) given (\tilde{a}, \tilde{b}) . Theorem 2.4 of (Mangasarian and Shiau, 1987) proves that \mathcal{M} is a lower-semicontinuous map (image of linear inequalities under perturbations in the right-hand side). Michael's Theorem (9.1.2 in (Aubin and Frankowska, 2009)) suffices to show that a continuous selection exists, because $\mathcal{M}_{\gamma, C}$ is lower-semicontinuous with closed convex images, \mathcal{P} is compact, and $\mathbb{R}^{n_{cl}}$ is a Banach space. \square

Theorem 3.2. The bounds from Eq. (16) will converge to $\lim_{d \rightarrow \infty} \gamma_d^* = \gamma^*$ from Problem 1 with $\gamma_d^* \geq \gamma_{d+1}^* \geq \dots \gamma^*$ when $\bar{\mathcal{P}}$ is Archimedean.

Proof. Let (γ, m) be a solution to Problem 1 with each $m_i(a, b)$ continuous over \mathcal{P} (by Theorem 3.1). For each $\epsilon > 0$, there exists functions $\tilde{m}_i(a, b) \in \mathbb{R}[a, b]$ such that $\sup_{(a,b) \in \mathcal{P}} |(m_i(a, b) + \epsilon) - \tilde{m}_i(a, b)| \leq \epsilon$ by the Stone Weierstrass theorem over the compact \mathcal{P} . Define $r_i(a, b) = m_i(a, b) + \epsilon - \tilde{m}_i(a, b)$ as the approximation error. The finite-degree polynomial variates $(\gamma + n_{cl}\epsilon, \tilde{m}_i)$ are feasible solutions to (14) with

$$0 \leq m_i(a, b) - a_{cl}(a, b) \leq \tilde{m}_i(a, b) - a_{cl}(a, b) \quad (17a)$$

$$0 \leq \gamma - \sum_i m_i(a, b) \leq (\gamma + n_{cl}\epsilon) - \sum_i \tilde{m}_i(a, b). \quad (17b)$$

The multipliers for Putinar Psatz certificates for (17a) and (17b) have finite (exponential) degree in terms of $\deg \tilde{m}_i$ (Nie and Schweighofer, 2007). Therefore it holds that for each ϵ , there exists some degree d such that $\gamma_d = \gamma + n_{cl}\epsilon$. Because $\lim_{\epsilon \rightarrow 0} \gamma^* + n_{cl}\epsilon = \gamma^*$, it holds that $\lim_{d \rightarrow \infty} \gamma_d^* = \gamma^*$. The sequence is decreasing with $\gamma_d^* \geq \gamma_{d+1}^*$ because the WSOS cones satisfy $\Sigma[\mathcal{P}]_{\leq 2d} \subset \Sigma[\mathcal{P}]_{\leq 2(d+1)}$. \square

4 Alternatives Control

Eq. (16) involves a total of $2(n_a + n_b + T) - 1$ variables $(a, b, \Delta u, \Delta y)$. The Psatz expressions such as in (16d) will therefore have Gram matrices of size $\binom{2(n_a + n_b + T) + d}{d}$ at each fixed degree d . Performance of Eq. (16) is therefore polynomial in T for fixed (n_a, n_b, d) , and is jointly combinatorial in all parameters.

A theorem of alternatives may be utilized to eliminate the noise terms $(\Delta u, \Delta y)$ from Putinar expressions. This Alternatives algorithm scales in a linear manner based on T and possesses Gram matrices of maximal size $\binom{n_a + n_b + d}{d}$. Letting $n_a = 3, n_b = 2, T = 10$, this maximal size is $\binom{15}{2} = 465$ for Full and $\binom{5+1}{1} = 6$ for Alternatives.

4.1 ARX Alternatives Psatz

Let $q(a, b)$ be a function defined over \mathcal{P} . Because $q(a, b)$ is a function of (a, b) alone, the following positivity criteria are equivalent by projection:

$$q(a, b) > 0 \quad \forall (a, b) \in \mathcal{P} \quad (18a)$$

$$q(a, b) > 0 \quad \forall (a, b, \Delta u, \Delta y) \in \bar{\mathcal{P}}. \quad (18b)$$

The following statement is a strong alternative to (18):

$$\exists (a, b, \Delta u, \Delta y) \in \bar{\mathcal{P}} : \quad -q(a, b) \geq 0. \quad (19)$$

Dual variable functions may be defined according to the Putinar multipliers in constraint description (12) with

$$\psi_t^\pm(a, b) \geq 0 \quad \forall t = (-n_b + 1)..T - 1 \quad (20a)$$

$$\zeta_t^\pm(a, b) \geq 0 \quad \forall t = (-n_a + 1)..T \quad (20b)$$

$$\mu_t(a, b) \quad \forall t = 1..T. \quad (20c)$$

The (a, b) dependence in terms $(\psi^\pm, \zeta^\pm, \mu)$ will be omitted to simplify notation. Additionally, the term ψ^+ will refer to the vector $\{\psi_t^+\}_{t=1}^T$ (with similar vectorial definitions for ψ^-, ζ^\pm, μ). The nonnegativity constraints in (20a) and (20b) are required to hold for all $(a, b) \in \mathcal{P}$.

A weighted sum Φ may be developed from q , multipliers in (20), and the description (12), by forming

$$\begin{aligned} \Phi = & -q(a, b) + \sum_{t=1}^T \mu_t h_t \\ & + \sum_{t=1}^T \mu_t (\sum_{i=1}^{n_b} b_i \Delta u_{t-i} - \sum_{i=1}^{n_a} a_i \Delta y_{t-i} - \Delta y_t) \\ & + \sum_{t=-(n_b-1)}^{T-1} (\psi_t^+ (\epsilon_u - \Delta u_t) + \psi_t^- (\epsilon_u + \Delta u_t)) \\ & + \sum_{t=-(n_a-1)}^T (\zeta_t^+ (\epsilon_y - \Delta y_t) + \zeta_t^- (\epsilon_y + \Delta y_t)). \end{aligned} \quad (21)$$

The terms in (21) that are constant in $(\Delta u, \Delta y)$ may be collected into

$$\begin{aligned} Q(a, b) &= -q(a, b) + \epsilon_u \mathbf{1}^T (\psi^+ + \psi^-) \\ &\quad + h^T \mu + \epsilon_y \mathbf{1}^T (\zeta^+ + \zeta^-). \end{aligned} \quad (22)$$

Using the cross-correlation operator \star from (1), Eq. (21) may be rewritten as

$$\begin{aligned} \Phi &= Q(a, b) + (\mu \star b)^T \Delta u - (\mu \star [1; a])^T \Delta y \\ &\quad + (\psi^- - \psi^+)^T \Delta u + (\zeta^- - \zeta^+)^T \Delta y. \end{aligned} \quad (23)$$

A Lagrangian dual function $g(a, b)$ may be defined as

$$g(a, b) = \sup_{\Delta u \in \mathbb{R}^{T+n_b}, \Delta y \in \mathbb{R}^{T+n_a}} \Phi(a, b, \Delta u, \Delta y). \quad (24)$$

The coefficients of $(\Delta u, \Delta y)$ in (23) must be zero in order to ensure that $g(a, b)$ is bounded. The value of this dual function is

$$g(a, b) = \begin{cases} Q(a, b) & \psi^+ - \psi^- = \mu \star b \\ & \zeta^+ - \zeta^- = \mu \star [1; a] \\ \infty & \text{Else.} \end{cases} \quad (25)$$

Problem 2. A feasibility program to certify (18) is

$$\underset{\psi^\pm, \zeta^\pm, \mu \text{ from (20)}}{\text{find}} \quad -Q(a, b) > 0 \quad \forall (a, b) \in \mathcal{P} \quad (26a)$$

$$\psi^+ - \psi^- = \mu \star b \quad (26b)$$

$$\zeta^+ - \zeta^- = \mu \star [1; a]. \quad (26c)$$

Theorem 4.1. Problem 2 certifies (18) and is a strong alternative of (19).

Proof. Sufficiency: The statement $g(a, b) < 0$ is a sufficient condition for invalidation of (19). Assuming that (19) holds with $-q \geq 0$, then Φ is constructed by the addition of the nonnegative $-q$ plus nonnegative weights (ϕ^\pm, ζ^\pm) times nonnegative constraints (ϵ_u, ϵ_y) plus free weights μ_t times data-consistency equality conditions (11). The supremal value of a nonnegative term Φ being negative with $g(a, b) < 0$ is a contradiction.

Necessity: The constraints in (12) are affine $(\Delta u, \Delta y)$. Necessity follows if constraints are concave (including affine) in the eliminated variables by Section 5.8 (Boyd et al., 2004). \square

Proposition 4.2. The multiplier functions $(\psi^\pm, \zeta^\pm, \mu)$ may each be chosen to be polynomial in the compact set \mathcal{P} .

Proof. This proof is omitted for brevity, as it follows using arguments from the proof of Theorems 4.4 (continuity) and 4.5 (polynomial approximability) of (Miller et al., 2022a). \square

Forming a Psatz with Problem 2 requires an additional assumption:

A2 An Archimedean set $\Pi \supseteq \mathcal{P}$ is known in advance.

Eq. (27) is a Psatz that can certify (18) at degree d .

$$\psi^\pm(a, b) \in (\Sigma[\Pi]_{\leq 2d})^{T+n_b-1} \quad (27a)$$

$$\zeta^\pm(a, b) \in (\Sigma[\Pi]_{\leq 2d})^{T+n_a} \quad (27b)$$

$$\mu(a, b) \in (\mathbb{R}[a, b]_{\leq 2d-1})^T \quad (27c)$$

$$-Q(a, b; \psi^\pm, \zeta^\pm; \mu) \in \Sigma[\Pi]_{\leq 2d} \text{ (from (22))} \quad (27d)$$

$$\psi^+ - \psi^- = \mu \star b, \quad \zeta^+ - \zeta^- = \mu \star [1; a]. \quad (27e)$$

Proposition 4.3. Eq. (27) will converge to a positivity certificate for (18) as $d \rightarrow \infty$ under A1 and A2.

Proof. Proposition 4.2 ensures that there exists at least one polynomial certificate $(\psi^\pm, \zeta^\pm, \mu)$. Given that polynomials have finite degree, letting $d \rightarrow \infty$ will ensure that a polynomial will be reached at some finite d . \square

4.2 Alternatives for Superstabilization

The Alternatives Psatz in Eq. (27) may be employed for superstabilizing EIV control. The Alternatives program to solve Problem 1 is described in Eq. (28). Eq. (28) is structurally identical to (16), with the difference that m is now a function of (a, b) and the positivity constraints in (28d)-(28f) are imposed using Eq. (27) ($\Sigma[\mathcal{P}]_{\leq 2d}^{\text{alt}}$).

$$\gamma_d^* = \min_{\gamma \in \mathbb{R}} \gamma \quad (28a)$$

$$\tilde{a} \in \mathbb{R}^{n_a}, \tilde{b} \in \mathbb{R}^{n_b} \quad (28b)$$

$$m_i \in (\mathbb{R}[a, b])_{\leq 2d}^{n_{cl}} \quad \forall i \in 1..n_{cl} \quad (28c)$$

$$\gamma - \sum_{i=1}^{n_{cl}} m_i(a, b) \in \Sigma[\tilde{\mathcal{P}}]_{\leq 2d} \quad (28d)$$

$$m_i(a, b) - a_i^{cl}(a, b) \in \Sigma[\tilde{\mathcal{P}}]_{\leq 2d} \quad \forall i = 1..n_{cl} \quad (28e)$$

$$m_i(a, b) + a_i^{cl}(a, b) \in \Sigma[\tilde{\mathcal{P}}]_{\leq 2d} \quad \forall i = 1..n_{cl}. \quad (28f)$$

5 Computational Complexity

This section will tabulate the computational complexity of imposing that a polynomial $q(a, b) \in \mathbb{R}[a, b]_{\leq 2d}$ is nonnegative over \mathcal{P} using the Putinar (Full) Psatz in (7) and the Alternatives Psatz in (27). We will use the quantity $N = n_a + n_b$ in this analysis. The Full program in (16) and the Alternatives program in Eq. (28) each have $2n_{cl} + 1$ instances of their respective Psatz certificates. Table 1 compiles the sizes of the optimization variables in the Full and Alternatives programs. The notation \mathbb{R} and \mathbb{S}_+ in the table refers to the length of a real vector and the dimension of a PSD matrix, respectively. This analysis treats the set $\Pi = \mathbb{R}^{n_a+n_b}$ in the Alternatives program to simplify tabulation. If the BSA set Π has N_Π polynomial-defined constraints, then the Alternatives program has 1 set of variables corresponding to entries in the Alternatives column in 1 and N_Π sets of variable with smaller sizes.

Table 1: Size of Superstabilizing Psatz.

	# polynomials	Full	Alternatives
σ_0	1	$\mathbb{S}_+ \binom{2(N+T)-1+d}{d}$	$\mathbb{S}_+ \binom{N+d}{d}$
ψ^\pm	$2(n_b + T - 1)$	$\mathbb{S}_+ \binom{2(N+T)-1+d-1}{d-1}$	$\mathbb{S}_+ \binom{N+d}{d}$
ζ^\pm	$2(n_a + T)$	$\mathbb{S}_+ \binom{2(N+T)-1+d-1}{d-1}$	$\mathbb{S}_+ \binom{N+d}{d}$
μ	T	$\mathbb{R} \binom{2(N+T)-1+2d-2}{2d-2}$	$\mathbb{R} \binom{N+2d-1}{2d-1}$

Remark 3. *The Alternatives program is more efficient than the Full program for each d given that $N < 2(N + T)$.*

Remark 4. *Section 4 eliminated $(\Delta u, \Delta y)$ and presented a Psatz (27) in terms of the $(n_a + n_b)$ variables (a, b) . Given that $\tilde{\mathcal{P}}$ in (12) is bilinear in terms of the groups $[(a, b), (\Delta u, \Delta y)]$ and each a^{cl} in (5) is linear in (a, b) , an Alternatives program in terms of $(\Delta u, \Delta y)$ could have been created by eliminating (a, b) . This approach would be more complex than Eq. (28), because $(\Delta u, \Delta y)$ has a total of $2T + n_a + n_b - 1 > n_a + n_b$ variables.*

6 Numerical Examples

MATLAB (2020b) code to reproduce the below experiments is located at https://github.com/jarmill/eiv_arx. These routines require Mosek (ApS, 2020) and YALMIP (Lofberg, 2004).

We tested the effectiveness of the proposed method using a discrete-time model of

$$G(\lambda) = \frac{\lambda^2}{1 + 0.5\lambda - 1.21\lambda^2 - 0.605\lambda^3}. \quad (29)$$

This system is open-loop unstable with unstable poles $z = \frac{1}{\lambda} = \pm 1.1$. For comparison purpose, we first solve the model-based superstabilization problem. This is addressed by minimizing γ with $\|a^{cl}\|_1 \leq \gamma$ and known a, b from (29) to search for control coefficients \tilde{a}, \tilde{b} . A superstabilizing controller will occur with $\gamma < 1$, and a smaller γ results in a faster controller. A special case $\gamma = 0$ corresponds to the deadbeat control, i.e. all closed-loop poles are located at the origin. For model (29), two types of controllers can be found. The low-order controller is obtained with $n_{\tilde{a}} = 3, n_{\tilde{b}} = 2$ and $\gamma = 0.4417$:

$$C(\lambda) = \frac{1.829\lambda^2}{1 - 0.5\lambda + 1.46\lambda^2 - 0.73\lambda^3}. \quad (30)$$

The deadbeat controller is obtained with $n_{\tilde{a}} = 4, n_{\tilde{b}} = 3$ and $\gamma = 0$:

$$C(\lambda) = \frac{0.73\lambda - 1.464\lambda^2 - 0.8833\lambda^3}{1 - 0.5\lambda + 1.46\lambda^2} \quad (31)$$

Remark 5. Note that we design using $\lambda = \frac{1}{z}$, the controller is improper in λ but is physically realizable in z .

For the data-driven setup, we used the deadbeat controller as the benchmark, i.e. $\gamma = 0$. The system is excited using uniformly distributed input $\|u_t\|_\infty \leq 1$, initial output $\{y_t\}_{t=-na+1}^{-1}$ and noise $\|\Delta y\|_\infty \leq \epsilon_y, \|\Delta u\|_\infty \leq \epsilon_u$ with $T = 10$ samples. We start from the noise-free data, i.e. $\epsilon_y = \epsilon_u = 0$. Directly applying Full Eq. (16) (with $d = 2$) leads to an intractable problem. Using the Alternatives Eq. (28) (with $d = 1$) leads to the same ℓ_1 -optimal deadbeat controller as in (31). This indicates that there is no gap between the original problem and its alternative form. For specific complexity, see Table 2.

Table 2: Size of Superstabilizing Psatz (Model (29) with $n_a = 3, n_b = 2, T = 10$).

	# polynomials	Full	Altern.
σ_0	1	465	6
ψ^\pm	22	30	6
ζ^\pm	26	30	6
μ	10	465	6

Next, we show the effect of changing T (Table 3) and ϵ (Table 4) on the performance index γ .

Table 3: γ v.s. T with $\epsilon_y = \epsilon_u = 0.02$.

T	20	40	60	80
γ	0.4365	0.3132	0.2732	0.2515

Table 4: γ v.s. $\epsilon = \epsilon_y = \epsilon_u$ with $T = 80$.

ϵ	0.02	0.04	0.06	0.08
γ	0.2515	0.4924	0.7312	0.9755

Several conclusions can be drawn from the tables.

- a) For the noisy trajectory, $\gamma \neq 0$, the learned controller is no longer deadbeat.
- b) As we increase T , the consistency set shrinks, hence we get a faster controller.
- c) Larger noise leads to a larger consistency set, which makes it harder to find a single robust controller. This issue may be alleviated by collecting more samples.

The performance γ can also be improved by selecting a higher order controller. This is illustrated by Table 5 in which $n_b = n_a - 1$.

Remark 6. This behaviour is not seen in the model-based control since any controller order with $n_a \geq 4, n_b \geq 3$ leads to the deadbeat control. However, for data-driven control, the extra order helps to reduce γ .

Table 5: γ vs. n_a with $\epsilon_y = \epsilon_u = 0.01, T = 10$.

n_a	4	6	8	10
γ	0.6926	0.5436	0.5167	0.5166

7 Extensions

7.1 Process Noise

L_∞ -bounded process noise w_t with $\|w\|_\infty \leq \epsilon_w$ can be added into the ARX model (2) with

$$y_t = -\sum_{i=1}^{n_a} a_i y_{t-i} + \sum_{i=1}^{n_b} b_i u_{t-i} + w_t. \quad (32)$$

With process noise, the equality constraint in (11) is relaxed to a pair of inequalities for each $t = 1..T-1$ as

$$-\epsilon_w \leq h_t + \left(-\sum_{i=1}^{n_a} a_i \Delta y_{t-i} + \sum_{i=1}^{n_b} b_i \Delta u_{t-i}\right) - \Delta y_t \leq \epsilon_w. \quad (33)$$

The full program in (16) may be executed with process noise by posing Psatz constraints over an enlarged consistency set of

$$\bar{\mathcal{P}}_w : \left\{ \begin{array}{l|l} a \in \mathbb{R}^{n_a}, b \in \mathbb{R}^{n_b} & \|\Delta u\|_\infty \leq \epsilon_u \\ \Delta u \in \mathbb{R}^{T+n_b-1} & \|\Delta y\|_\infty \leq \epsilon_y \\ \Delta y \in \mathbb{R}^{T+n_a} & \text{Ineq. (33) holds } \forall t = 1..T \end{array} \right\} \quad (34)$$

The Alternatives Psatz in (27) may be derived for the set $\bar{\mathcal{P}}_w$ by defining functions $\mu_t^\pm(a, b)$ that are nonnegative over $\pi^{a,b} \bar{\mathcal{P}}_w$ replacing (22) with

$$\begin{aligned} Q^w(a, b) = & -q(a, b) + h^T(\mu^+ - \mu^-) + \epsilon_w \mathbf{1}^T(\mu^+ + \mu^-) \\ & + \epsilon_u \mathbf{1}^T(\psi^+ + \psi^-) + \epsilon_y \mathbf{1}^T(\zeta^+ + \zeta^-), \end{aligned} \quad (35)$$

and every instance of μ in (27c)-(27c) with $\mu^+ - \mu^-$.

7.2 Switched ARX Models

The switched setting involves N_s ARX models (subsystems) with per-subsystem parameters (a^k, b^k) and lengths (n_a^k, n_b^k) for $k = 1..N_s$. A switching sequence S takes on values $S(t) \in 1..N_s$ for each $t \in 1..T$. The ground-truth switched ARX model given S is

$$y_t = -\sum_{i=1}^{n_a^{S(t)}} a_i^{S(t)} y_{t-i} + \sum_{i=1}^{n_b^{S(t)}} b_i^{S(t)} u_{t-i}. \quad (36)$$

Performing a substitution as in (8) allows for the creation of a consistency set $\bar{\mathcal{P}}^S(\{a_k, b_k\}_{k=1}^{N_s}, \Delta u, \Delta y)$. The projection $\pi^{a,b} \bar{\mathcal{P}}^S(\{a_k, b_k\}_{k=1}^{N_s})$ has $n_{\text{sys}} = \sum_{k=1}^{N_s} n_a^k + n_b^k$ variables. The superstabilization problem for switched ARX with a known switching sequence involves finding subsystem-controllers $(\tilde{a}^k, \tilde{b}^k)$ such that each subsystem-closed-loop transfer functions in (4) is superstable. An Alternatives Psatz (modification of (27)) would involve Gram matrices of maximal size $\binom{n_{\text{sys}}+d}{d}$.

8 Conclusion

This paper proposed an algorithm to perform superstabilizing control of EIV-corrupted ARX systems. The controller (\tilde{a}, \tilde{b}) is recovered by solving a POP using the moment-SOS hierarchy of SDPs. Utilizing the Theorem of Alternatives results in a significantly more tractable program as compared to the Full case.

Future work involves forming superstabilizing controllers for Multi-Input Multi-Output systems using Matrix Fraction Descriptions (Chen et al., 2005) and for Linear-Parameter Varying systems.

References

- M. ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.*, 2020. URL <https://docs.mosek.com/9.2/toolbox/index.html>.
- J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Springer Science & Business Media, 2009.
- J. Berberich, A. Koch, C. W. Scherer, and F. Allgöwer. Robust data-driven state-feedback design. In *2020 American Control Conference (ACC)*, pages 1532–1538. IEEE, 2020.
- F. Blanchini and M. Sznaier. Persistent disturbance rejection via static-state feedback. *IEEE Transactions on Automatic Control*, 40(6):1127–1131, 1995.
- S. Boyd, S. P. Boyd, and L. Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.
- V. Cerone. Feasible parameter set for linear models with bounded errors in all variables. *Automatica*, 29(6):1551–1555, 1993.
- V. Cerone, D. Piga, and D. Regruto. Set-Membership Error-in-Variables Identification Through Convex Relaxation Techniques. *IEEE Transactions on Automatic Control*, 57(2):517–522, 2011.
- J. Chen, C. M. Lagoa, and A. Ray. On the design of discrete-time fixed-order controllers for persistent disturbance rejection. *IFAC Proceedings Volumes*, 38(1):360–365, 2005.
- Y. Cheng. *Robust Model Fitting via Convex Optimization Techniques*. Northeastern University, 2016.
- Y. Cheng, M. Sznaier, and C. Lagoa. Robust superstabilizing controller design from open-loop experimental input/output data. *IFAC-PapersOnLine*, 48(28):1337–1342, 2015. 17th IFAC Symposium on System Identification SYSID 2015.
- T. Dai and M. Sznaier. A moments based approach to designing mimo data driven controllers for switched systems. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 5652–5657. IEEE, 2018.
- C. De Persis and P. Tesi. Formulas for Data-Driven Control: Stabilization, Optimality, and Robustness. *IEEE Trans. Automat. Contr.*, 65(3):909–924, 2020. doi: 10.1109/TAC.2019.2959924.
- J. Gouveia, P. A. Parrilo, and R. R. Thomas. Lifts of Convex Sets and Cone Factorizations. *Mathematics of Operations Research*, 38(2):248–264, 2013.
- Z.-S. Hou and Z. Wang. From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, 235:3–35, 2013. ISSN 0020-0255. doi: <https://doi.org/10.1016/j.ins.2012.07.014>. Data-based Control, Decision, Scheduling and Fault Diagnostics.
- J. B. Lasserre. *Moments, Positive Polynomials And Their Applications*. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009. ISBN 9781908978271.
- J. Lofberg. Yalmip : a toolbox for modeling and optimization in matlab. In *ICRA (IEEE Cat. No.04CH37508)*, pages 284–289, 2004.
- O. L. Mangasarian and T.-H. Shiau. Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems. *SIAM Journal on Control and Optimization*, 25(3):583–595, 1987.
- J. Miller, T. Dai, and M. Sznaier. Data-Driven Stabilizing and Robust Control of Discrete-Time Linear Systems with Error in Variables, 2022a. arXiv:2210.13430.
- J. Miller, T. Dai, and M. Sznaier. Data-Driven Superstabilizing Control of Error-in-Variables Discrete-Time Linear Systems. In *2022 61st IEEE Conference on Decision and Control (CDC)*, pages 4924–4929, 2022b.
- J. Nie and M. Schweighofer. On the complexity of putinar’s positivstellensatz. *Journal of Complexity*, 23(1):135–150, 2007.

- J. Norton. Identification of parameter bounds for ARMAX models from records with bounded noise. *International Journal of control*, 45(2):375–390, 1987.
- K. Ogata. *Discrete-time control systems*. Prentice-Hall, Inc., 1995.
- B. Polyak and M. Halpern. Optimal design for discrete-time linear systems via new performance index. *International Journal of Adaptive Control and Signal Processing*, 15(2):129–152, 2001.
- B. T. Polyak and P. S. Shcherbakov. Superstable Linear Control Systems. I. Analysis. *Automation and Remote Control*, 63(8):1239–1254, 2002.
- M. Putinar. Positive Polynomials on Compact Semi-algebraic Sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993. ISSN 00222518, 19435258.
- T. Söderström. *Errors-in-Variables Methods in System Identification*. Springer, 2018.
- M. Sznaier and F. Bianchini. Mixed l_∞/h_∞ suboptimal controllers for siso continuous-time systems. *IEEE Transactions on Automatic Control*, 40(11):1831–1840, 1995. doi: 10.1109/TAC.1995.7469037.
- H. J. van Waarde, M. K. Camlibel, and M. Mesbahi. From noisy data to feedback controllers: non-conservative design via a matrix S-lemma. *IEEE Trans. Automat. Contr.*, 2020.