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

Confidently incorrect: nonlinear observers with online error bounds

Jonathan Bunton and Paulo Tabuada

Abstract— **Feedback control typically relies on an estimate of the system state provided by an estimation scheme. These estimates, however, are always affected by errors that have non-negligible impacts on control performance. Various stabilizing and safety-critical control frameworks address this issue, but all require some characterization of the current estimation error to determine when to apply more or less conservative control inputs. Current methods of bounding these errors either take a very coarse worst-case bound or employ computationally expensive time-varying set-valued methods.**

This paper fills the missing gap in these works, presenting new deterministic worst-case error bounds for a state estimation scheme for generic nonlinear systems. Crucially, these error bounds can be efficiently computed in real-time and shrink or grow depending on the current system behavior and the current measurement quality. These new, lightweight, "online" error bounds can directly interface with the aforementioned measurement-robust control frameworks, resulting in less conservative control actions while retaining safety and stability guarantees.

I. Introduction

In feedback control, one typically builds a full or partial-state feedback control law to accomplish the desired control task. This is particularly true in safety-critical scenarios, where one must prioritize the system's safety above all else. Almost all of these techniques for nonlinear control–particularly in safety-critical control–rely on knowledge of the system state. In practice, this means that the state feedback control law is designed first, and then implemented using, not the true system state, but an estimate from a separately designed state estimator.

While theoretically justified in some cases, the choice of estimation scheme can have major impacts on the overall control performance. It is well-known, for example, that stabilizing control laws for nonlinear systems may catastrophically fail when instead given a state *estimate* [1].

Many modern nonlinear control techniques have been developed that accommodate the inherent imperfect knowledge of the state in a measurement-robust or uncertainty-aware framework. For example, the robust control Lyapunov and Barrier function frameworks have both been adapted to handle uncertainty in estimation [2], [3], [4]. These frameworks, however, must assume some bound on the state estimation error and employ more conservative control actions depending on the magnitude of the error bound. In safety-critical control, for example, the measurement-robust barrier function framework effectively "inflates" the unsafe set and attempts to maintain a harsher safety criterion [3]. Beyond conservative control inputs, loose bounds can also lead to issues where a guaranteed safe control input does not exist.

What these existing measurement-robust frameworks lack is an estimation scheme that comes with error guarantees that *vary with time* in order to be less conservative, as noted in [4]. When equipped with such an estimation method, the measurement-robust control frameworks can adapt to be more or less conservative as the estimation error bounds grow or shrink. This adaptation may even be necessary in order to properly guarantee the safety or stability of the closed-loop system.

Many nonlinear estimation methods–even classic algorithms such as the Extended Kalman Filter (EKF)–already have some form of time-varying error guarantee [1]. When these guarantees exist, however, they typically include a fixed inflation to accommodate the *worst-case* measurement noise or disturbances in the system, in a manner akin to input-to-state stability (ISS) bounds [5]. These bounds are then always "inflated" regardless of the *actually experienced* measurement noise or disturbances, even if the observer itself may be performing better in some periods than others.

Alternatively, there exist *set-valued observers* that hold on to *tight, time-varying* error bounds that may shrink or grow depending on the exact sequence of system outputs. Set-valued observers, however, are typically only available for highly structured or linear systems [6]. Even when available, these methods are often extremely computationally demanding, limiting their practical utility [7].

In this work, we present an estimation scheme based on numerical differentiation that directly targets these issues: it possesses deterministic, time-varying bounds that adapt online to the *experienced* measurement noise and system behavior. These new guarantees can directly be handed to any measurement-robust control framework, where their time-varying nature permits more aggressive control actions when the estimation method is more confident. Moreover, since these guarantees are deterministic worst-case bounds, any measurement-robust control law based on these values will yield deterministic worst-case correctness proofs.

II. Problem setup and background

We consider nonlinear control systems of the form:

$$\begin{aligned} \dot{x} &= f(x,u) \\ y &= h(x,u), \end{aligned} \tag{1}$$

This work was partially supported by NSF award 2211146 and 2146828.

J. Bunton and P. Tabuada are with the UCLA Electrical and Computer Engineering Department, Los Angeles, CA 90095. {j.bunton,tabuada}@ucla.edu

979-8-3503-8265-5/$31.00 ©2024 AACC

where for all $t \in \mathbb{R}_{\geq 0}$, $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, and $y(t) \in \mathbb{R}^p$ is the output. We assume that the map f is sufficiently regular for solutions of (1) to exist and be unique for all $t \geq 0$ and $x(0) \in \mathbb{R}^n$.

We are interested in estimating the state $x(t)$ of the system (1) at some time $t \in [t_0, t_N]$ given (possibly noise-corrupted) $N+1$ sampled-data measurements of $y(t)$ at a window of times $\{t_0, t_1, ..., t_N\}$. The method we propose relies on the following definition of observability to ensure a well-posed problem.

Assumption 1: The control system (1) is *differentially observable of order* d. In particular, there exists a (possibly time-dependent) continuous function $\mathcal{L}$ that maps the d derivatives of the output $y(t)$ and input $u(t)$ to the state $x(t)$. More explicitly, $\mathcal{L}$ is such that:

$$(y(t), \dot{y}(t), ..., y^{(d)}, u(t), \dot{u}(t), ..., u^{(d)}(t)) \xmapsto{\mathcal{L}} x(t).$$

As described in [8], [9], places where the map $\mathcal{L}$ fails to exist are called *singular* observations.

A. Related work

A number of frameworks have addressed the interface between uncertainty in the system state and control in the context of stability and safety. We only discuss non-stochastic methods here, as they are closer to our work and its deterministic guarantees. From the perspective of stability, there are characterizations of ISS with respect to estimation errors, which guarantee that bounded errors cannot unboundedly destroy stability [10]. In practice, however, there is no method for creating controllers that enforce this particular form of ISS for arbitrary nonlinear systems [11].

Other works developed notions of robustness to estimation errors, leading to the concepts of *robust* control Lyapunov functions (CLFs) and measurement-robust control barrier functions (CBFs) [2], [3], [4]. Both of these schools of thought, however, rely on a characterization of the estimation error that is valid at any instant of time. Loose offline characterizations of this uncertainty can lead to overly conservative controls, or worse, issues of feasibility from a lack of "guaranteed safe/stable" control actions.

On the estimation side of this problem, there are many methods for state estimation that possess a time-varying error bound. Perhaps the most straightforward to understand are set-valued observers, wherein a tight approximation (polyhedral, ellipsoidal, or a hyper-rectangle) of the possible states of the system is propagated through the dynamics at each step [7]. This tighter representation is less conservative, but comes at the cost of limited applicability and often prohibitively large amounts of computation and memory.

More familiar observers possess asymptotic guarantees, and even ISS-like guarantees are often established [5]. Even these ISS guarantees, however, rely on some *a priori estimate* of the worst-case measurement noise for all time, then inflate the estimation guarantees *for all time* accordingly. Some promising and recent notions of ISS observers with "fading memory" exist through the use of input-to-state *dynamical* stability, but the error bounds provided by these estimators are not always available in real-time to mitigate the issues in measurement-robust control [12].

While developing observers for systems such as (1), one natural thought is to directly leverage Assumption 1 and consider derivative estimation equivalent to state estimation [8]. This idea is not new, and many existing works connected numerical differentiation techniques to state estimation dating back to the 1990s, proving that these estimation techniques can produce globally bounded error [13]. The existing guarantees, however, are strictly *offline*: given some estimate of the output's nonlinearity and magnitude of the noise, a single static error bound is provided for all time.

In this work, we show that these offline guarantees can be *significantly tightened* and made *online*. In particular, we prove a time-varying estimation error bound for Savitzky-Golay filtering that can be computed online with a simple multiplication by a fixed pre-computed matrix. Moreover, we show in experiments that these online bounds are *orders of magnitude tighter* than previously established offline bounds.

B. Savitzky-Golay filtering

The differential observability condition effectively equates estimating the *state* of the control system with estimating its output and derivatives. As such, we will construct a method for estimating d derivatives of the output from sampled-data measurements that possesses the online error bounds we seek.

We propose a state estimation framework built on a classical scheme for numerical differentiation: polynomial least-squares, or *Savitzky-Golay filtering* [14].

In Savitzky-Golay filtering, we build a local (in time) approximation of the output signal y by fitting a window of $N+1$ samples in some interval $[t_i, t_f] \subseteq \mathbb{R}$ with a degree-d polynomial. We then estimate the d derivatives of y at some time in this window $\tau \in [t_i, t_f]$ by differentiating the polynomial approximation to y at τ. Finally, we apply the map $\mathcal{L}$ from Assumption 1 to produce a state estimate $\hat{x}(\tau)$. Notably, if the samples are uniformly spaced, this entire process becomes a single matrix multiplication with a fixed matrix computed offline.

In this work, we appeal to the following intuition: the *residuals* from the least-squares regression naturally measure fit quality. By residuals, we mean the misfit between the polynomial p and the output y at the sampled outputs. These residuals may be high or low depending on the *actual* measurement noise and nonlinearities at any given time, rather than being fixed a priori. Our main results formalize this intuition by connecting the online residuals to a deterministic worst-case error bound on the derivative estimation error during Savitzky-Golay filtering.

III. Online error bounds

We assume that the output function h for the control system (1) is such that the output y is continuous and $d+1$-times differentiable. For simplicity of analysis, we discuss only scalar outputs ($m = 1$), as the generalization to higher dimensions is straightforward.

We then approximate the output locally with a degree-d polynomial $p : \mathbb{R} \to \mathbb{R}$ of the form:

$$p(t) = a_0 + a_1 t + \cdots + a_{d-1} t^{d-1} + a_d t^d. \tag{2}$$

Given $N + 1$ measurements of the output y at times $\{t_0, t_1, ..., t_N\} \subseteq \mathbb{R}$, each corrupted by some noise signal $e(t_i)$, $t_i \in \{t_0, t_1, ..., t_N\}$, we determine the polynomial p by minimizing the squared error in the following optimization problem:

$$\underset{a:=(a_0,\ldots,a_d)\in\mathbb{R}^{d+1}}{\text{minimize}} \quad \sum_{i=0}^{N} \|[y(t_i) + e(t_i)] - p(t_i)\|_2^2. \tag{3}$$

Note that we do not have access to $y(t_i)$, only its noisy measurements $y(t_i) + e(t_i)$ at each sampled time.

A. Error bounds on derivatives

First, we state our main result which holds with equality.

Theorem 1: Choose any subset of sample times $\mathcal{D} := \{s_0, s_1, ..., s_d\} \subseteq \{t_0, t_1, ..., t_N\}$ with cardinality $|\mathcal{D}| = d + 1 \leq N + 1$, and let $p : \mathbb{R} \to \mathbb{R}$ be any degree-d polynomial. Define the degree-d polynomial "residual interpolant" $r_{\mathcal{D}}$ associated with p and $\mathcal{D}$, i.e., the polynomial such that:

$$y(s_i) - p(s_i) = r_{\mathcal{D}}(s_i) \quad \text{for all } s_i \in \mathcal{D}. \tag{4}$$

Then for any $t \in [s_0, s_d]$, it holds that:

$$y^{(k)}(t) - p^{(k)}(t) = r_{\mathcal{D}}^{(k)}(t) + \frac{y^{(d+1)}(\xi)}{(d-k+1)!} \prod_{i=0}^{d-k} (t - \nu_i), \tag{5}$$

where $k \geq 0$ and $s_i \leq \nu_i \leq s_{i+k}$ for each $i = 0, 1, ..., d-k$ and $\xi \in [s_0, s_d]$.

Proof: Define the auxiliary function $Q : \mathbb{R} \to \mathbb{R}$ as:

$$Q(t) = y(t) - p(t) - r_{\mathcal{D}}(t).$$

By construction, Q is continuous and at least $d + 1$-times differentiable with at least $d+1$ zeroes in the interval $[s_0, s_d]$. In particular, its zeros are each of the $s_i \in \mathcal{D}$. By repeated applications of Rolle's Theorem, $Q^{(k)}$ has at least $d-k$ zeros, each denoted by ν_i, with $\nu_i \in [s_i, s_{i+k}]$.

Consider another function $H : \mathbb{R} \to \mathbb{R}$ defined as:

$$H(z) = Q^{(k)}(z) - \alpha \prod_{i=0}^{d-k} (z - \nu_i), \tag{6}$$

for some $\alpha \in \mathbb{R}$. Note that for any chosen $t \in [s_0, s_d]$ with $t \neq \nu_i$ for all $i = 0, 1, ..., d-k$, there exists a choice of $\alpha \in \mathbb{R}$ such that $H(t) = 0$. We will derive an explicit expression for this α in terms of $y^{(d+1)}$.

Because $H(t) = 0$ for $t \in [s_0, s_d]$, then H is $d - k + 1$ times differentiable with at least $d - k + 2$ zeros in the interval $[s_0, s_d]$. In particular, $H(z) = 0$ when $z = \nu_i$, with $i = 0, 1, ..., d-k$, and also at the prescribed $z = t$. Again using repeated applications of Rolle's Theorem, the $d-k+1$ derivative of H then has at least one zero in the interval $[s_0, s_d]$, meaning there exists some $\xi \in [s_0, s_d]$ (depending on t) such that:

$$\begin{aligned}
H^{(d-k+1)}(\xi) &= Q^{(k+(d-k+1))}(\xi) - \alpha(d-k+1)! \\
0 &= Q^{(d+1)}(\xi) - \alpha(d-k+1)! \\
&= y^{(d+1)}(\xi) - \alpha(d-k+1)! \\
\Rightarrow \ \alpha &= \frac{y^{(d+1)}(\xi)}{(d-k+1)!},
\end{aligned}$$

where in the third equality we abused the fact that p, $r_{\mathcal{D}}$, and $e_{\mathcal{D}}$ are degree-d polynomials.

Simply plugging this value for α into (6) and re-arranging, we find:

$$\begin{aligned}
H(t) = 0 &= Q^{(k)}(t) - \frac{y^{(d+1)}(\xi)}{(d-k+1)!} \prod_{i=0}^{d-k} (t - \nu_i) \\
&= y^{(k)} - p^{(k)} - r_{\mathcal{D}}^{(k)}(t) \\
&\quad - \frac{y^{(d+1)}(\xi)}{(d-k+1)!} \prod_{i=0}^{d-k} (t - \nu_i) \\
\Rightarrow \ y^{(k)} - p^{(k)} &= r_{\mathcal{D}}^{(k)}(t) + \frac{y^{(d+1)}(\xi)}{(d-k+1)!} \prod_{i=0}^{d-k} (t - \nu_i),
\end{aligned}$$

for all $t \in [s_0, s_d]$ as desired. ∎

Note that Theorem 1 is an *equality*, meaning there is no tighter bound for a given polynomial p. Our choice of polynomial p, however, will change the values (and derivatives) of the residual interpolant $r_{\mathcal{D}}$, suggesting we choose p that minimizes its impact (e.g., least-squares).

The equality (5) also behaves in expected ways for specific cases. If there is no measurement error $e(t_i) = 0$ and the function y *is* a polynomial of degree at most d, then the interpolating polynomial p has zero residuals, $y^{(d+1)}$ is uniformly zero, and therefore (5) guarantees zero misfit everywhere in the interval. Similarly, when the number of points and degree of the polynomial are equal ($d = N$), (5) immediately recovers the guarantee associated with *interpolating* polynomials.

Despite its tightness, Theorem 1 relies on knowledge of parameters that we do not have access to in reality: the noise-free values of $y^{(d+1)}(\xi)$, the times ν_i, and the underlying *true* misfit $y(t_i) - p(t_i)$. In practice, we only have access to the measured (noise-impacted) residuals, $y(t_i)+e(t_i)-p(t_i)$, and perhaps a uniform bound on the noise and value of $y^{(d+1)}(\xi)$. In the following corollary, we loosen the equality in (5) by only relying on these assumptions.

Corollary 1: Assume there exist $M, E \in \mathbb{R}_{\geq 0}$ such that $|y^{(d+1)}(\xi)| \leq M$ for all $\xi \in [s_0, s_d]$, and $|e(s_i)| \leq E$ for all $s_i \in \mathcal{D}$. If the subset $\mathcal{D}$ has maximal inter-sample spacing $s_{i+1} - s_i \leq \delta$, then:

$$\begin{aligned}
|y^{(k)}(t) - p^{(k)}(t)| \leq &\sum_{s_i \in \mathcal{D}} \left| l_i^{(k)}(t) \left(y(s_i) + e(s_i) - p(s_i)\right) \right| \\
&+ E \sum_{s_i \in \mathcal{D}} |l_i^{(k)}(t)| + M\delta^{d-k+1},
\end{aligned} \tag{7}$$

for all $k \geq 0$, where $l_i : \mathbb{R} \to \mathbb{R}$ with $i = 0, 1, ..., d$ are the Lagrange basis polynomials for $\mathcal{D}$:

$$l_i(t) = \prod_{s_j \in \mathcal{D} \setminus \{s_i\}} \frac{t - s_j}{s_i - s_j}. \tag{8}$$

Proof: We begin by simply applying the triangle inequality to the right-hand side of (5):

$$\begin{aligned} |y^{(k)}(t) - p^{(k)}(t)| &\leq |r_{\mathcal{D}}^{(k)}(t)| + \frac{|y^{(d+1)}(\xi)|}{(d-k+1)!} \prod_{i=0}^{d-k} |t - \nu_i| \\ &\leq |r_{\mathcal{D}}^{(k)}(t)| + \frac{M}{(d-k+1)!} \prod_{i=0}^{d-k} |t - \nu_i|. \end{aligned}$$

Then note that the interpolating polynomial $r_{\mathcal{D}}$ can be explicitly written as a function of its interpolation sites using the Lagrange basis for $\mathcal{D}$, as defined in (8):

$$\begin{aligned} &\left| \frac{d^k}{dt^k} r_{\mathcal{D}}(t) \right| = \left| \frac{d^k}{dt^k} \sum_{s_i \in \mathcal{D}} l_i(t) \left(y(s_i) - p(s_i) \right) \right| \\ &= \left| \sum_{s_i \in \mathcal{D}} l_i^{(k)}(t) \left(y(s_i) + e(s_i) - p(s_i) \right) - l_i^{(k)}(t) e(s_i) \right| \\ &\leq \sum_{s_i \in \mathcal{D}} \left| l_i^{(k)}(t) \left(y(s_i) + e(s_i) - p(s_i) \right) \right| + \left| l_i^{(k)}(t) e(s_i) \right| \\ &\leq \sum_{s_i \in \mathcal{D}} \left| l_i^{(k)}(t) \left(y(s_i) + e(s_i) - p(s_i) \right) \right| + E \left| l_i^{(k)}(t) \right|. \end{aligned}$$

Finally, we note that each of the ν_i is in the interval $[s_i, s_{i+k}]$. By assumption, each of these intervals is at most size $k\delta$. The product term is upper bounded by a choice of t that is at one end of the polynomial, which we can use as a lazy bound:

$$\begin{aligned} \frac{M}{(d-k+1)!} \prod_{i=0}^{d-k} |t - \nu_i| &\leq \frac{M}{(d-k+1)!} \prod_{i=0}^{d-k} (i+1) \cdot k\delta \\ &= M k^{d-k-1} \delta^{d-k+1}. \end{aligned}$$

While this bound is valid, it can *easily* be sharpened by characterizing this product for the specific choice of t where estimation is relevant. Combining these terms, we recover the desired result. ■

We have now removed any unknown quantities from Theorem 1, meaning Corollary 1 presents an *online-computable* bound characterizing the error in derivative estimation. Interestingly, this bound may vary in time with the fit *residuals* $y(t_i)+e(t_i)-p(t_i)$, which formalizes the intuition that "good polynomial fits" should produce better estimates, regardless of the standing assumptions on the system.

While the bounds in Corollary 1 are in principle "online computable", their practical value only holds if they are also computationally lightweight. Implementing both Savitzky-Golay filters *and evaluating Corollary 1's bounds* are computationally efficient. The filtering itself is a simple matrix multiplication of the current window of outputs by a fixed $N+1 \times N+1$ fitting matrix. The bounds require the measured residuals (one more matrix multiply and a vector subtraction) followed by a simple inner product with a (fixed, offline-computable) vector of $l_i^{(k)}(t)$ evaluations at the estimation time of interest $t \in [t_0, t_N]$.

Both Theorem 1 and Corollary 1 hold for an arbitrary degree-d polynomial p and its measured residuals. In practice, the Savitzky-Golay scheme uses the *least-squares* polynomial, which is useful because it indirectly minimizes the individual measured residuals in the bound (7).

One of the main motivations for using least-squares over interpolation is the ability to "smooth out" the impact measurement noise. This property is implicit in (7), where we can reduce the magnitude of the terms involving measurement noise by shrinking the values of the Lagrange basis polynomials $l_i^{(k)}(t)$ associated with the subset $\mathcal{D}$. To shrink these values, we must select a subset of times $\mathcal{D} \subseteq \{t_0, t_1, ..., t_N\}$ that is spaced as far apart as possible. If the polynomial p was selected with least-squares, then the term associated to its residuals maintains the same uniform bound regardless of the subset of fitting points $\mathcal{D}$. As we select a subset $\mathcal{D}$ with larger inter-sample times, however, the final term in (7) representing the output's deviation from polynomial grows. Choosing the best subset $\mathcal{D} \subseteq \{t_0, t_1, ..., t_N\}$ optimizes the trade off between *smoothing* and *accuracy*.

In principle, we could solve the combinatorial problem of choosing the subset $\mathcal{D} \subseteq \{t_0, t_1, ..., t_N\}$ with cardinality $|\mathcal{D}| = d+1$ that minimizes the bound (7) each time we make a derivative estimate. This approach is clearly intractable, but we can easily approximate it by choosing a small family of different subsets (e.g., by parameterizing the subset by several choices of inter-sample spacing δ) and evaluating (7) for each subset choice online, always claiming the tightest guarantee achieved by this family. Similarly, we could select the subset $\mathcal{D}$ *a priori* by assuming some fixed maximum values for the measured residuals and optimizing (7) over $\mathcal{D}$, but this reduces the dynamic properties of the bound.

Our method has two parameters: the degree of the fitting polynomial d, and the number of points in the window N. The bound in Corollary 1 scales with the fitting polynomial degree d in the exponent, suggesting d be as small as possible while still estimating enough derivatives. Similarly, as we increase the number of points in the window (assuming identical residual values), we increase the number of candidate subsets $\mathcal{D}$, therefore tightening the potential bound. These observations reflect typical rules for polynomial regression.

B. From derivatives to state

Up to this point, we have discussed only online error bounds for derivative estimation. We can transfer these error bounds from the space of derivative estimates to state estimates in a number of ways, such as interval analysis techniques. In practice, we may use whichever method provides the tightest guarantees, but for completeness, we state a naive but immediate result for the special case of Lipschitz continuous observability maps.

Theorem 2: Assume that the function $\mathcal{L}$ in Assumption 1 is uniformly continuous, and let $\hat{x}(t)$ denote the result of composing $\mathcal{L}$ with estimates of y and its d derivatives from a degree-d polynomial $p : \mathbb{R} \to \mathbb{R}$. Under the same setting as Corollary 1, there exists a nondecreasing function $\alpha : \mathbb{R} \to \mathbb{R}$ from the polynomial fit residuals to the estimation error:

$$|x(t) - \hat{x}(t)| \leq \alpha \left(\sum_{s_i \in \mathcal{D}} |y(s_i) + e(s_i) - p(s_i)| \right).$$

If an explicit expression for the observation map $\mathcal{L}$ is not known, we could follow the steps proposed in [15] and solve the observation equations (or the equations governing the derivatives) for the state via Newton's method. The error bounds would then propagate through the convergence guarantees of this method.

The process outlined above produces a state estimate for the time $t \in [t_0, t_N]$ where the derivative estimation takes place. Depending on when this time is chosen, there is necessarily a delay in the state estimate. We could choose to estimate derivatives at the most recent time $t_N \in [t_0, t_N]$, but differentiating fitting polynomials at their endpoints is notoriously inaccurate [16]. This difficulty is also reflected in the bounds given from Corollary 1, which are maximized when evaluated at t_0 and t_N.

We could also counteract the estimation delay by evolving the estimate from Theorem 2 forward with the differential equation model (1). We could even use an Extended Kalman Filter (EKF) initialized at the delayed estimate $\hat{x}(t)$ to both remove the delay and tighten the bounds from Theorem 2 when the EKF's local exponential convergence can be guaranteed [9], [13].

IV. Experiments and evaluation

In this section, we validate the theoretical bounds from Corollary 1 in a couple simple examples. In each case, we show that our online error bounds are orders of magnitude tighter than more standard offline bounds, and vary with time depending on the system dynamics.

A. Lorenz Attractor System

We consider the Lorenz attractor system dynamics with a single output:

$$\begin{aligned} \dot{x}_1 &= \sigma \cdot (x_2 - x_1) \\ \dot{x}_2 &= x_1 \cdot (\rho - x_3) - x_2 \ , \qquad y = x_1, \\ \dot{x}_3 &= x_1 x_2 - \beta x_3 \end{aligned}$$

where we set the parameters $\sigma = 10$, $\rho = 28$, and $\beta = \frac{8}{3}$. We use a sampling frequency of 100Hz (inter-sample time $\delta = 0.01$ seconds) and apply a Savitzky-Golay filter to fit a degree $d = 2$ polynomial to sliding windows of 20 measurements. We differentiate this polynomial at the midpoint, and in our comparisons we use the delayed value of the system output and state (meaning we do not consider the effects of estimation lag). We derive the error bounds on the state estimate by performing interval analysis on an explicit expression for the map from Assumption 1.

To highlight the *online* nature of our bounds, we add bounded ($E = 0.5$) measurement errors to the system *only during times* $t \in [1.6, 3.3]$, and otherwise we have zero noise. Crucially, we supply the bounds in Corollary 1 with the same value of $E = 0.5$ at all times, meaning we are *always* theoretically accommodating these measurement errors, even when none are present in the system. We also provide the bounds with a uniform bound $|y^{(d+1)}| \leq 96733$, which is valid for all time.

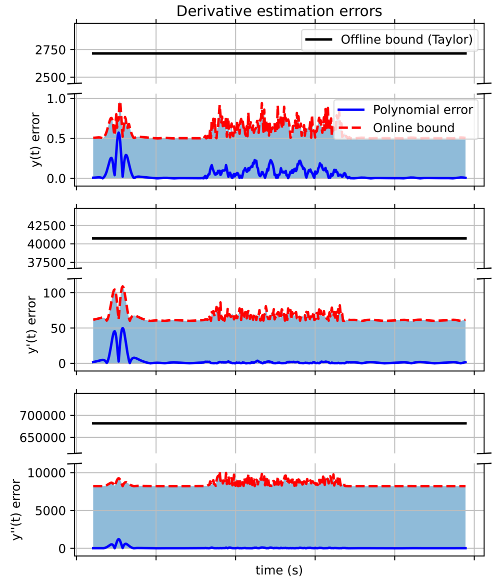

Fig. 1. Error in the the derivative estimation for the Lorenz system. The true estimation error is shown in blue, with dashed red lines and shading indicating the online error bounds of Corollary 1. The solid black lines denote offline bounds.

For comparison, we also plot some naive offline bounds derivable via Taylor series analysis, identical to those given in [13]. We omit the derivation of these bounds here for brevity, but the interested reader may find them in [17]. In Fig. 1, we show the error in the estimated output derivatives on a log-scale plot, highlighting that our new online bounds are *orders of magnitude tighter*. In addition to always being tighter, these bounds may adapt naturally the noise in the measurements. Furthermore, the measurement errors are bounded $E = 0.5$ and so the output's value (i.e., the estimate of the state x_1) cannot be more accurate than this fundamental limit. Similarly, the output's derivative estimation error is always lower bounded by $\frac{2E}{\delta_{max}} \approx 20$, and $\frac{4E}{\delta_{max}^2} \approx 800$ for the second derivative, where δ_{max} is the largest possible inter-sample spacing. Our error bounds in Fig. 1 show that our method is provably near these fundamental limits in the noiseless regime.

This same phenomenon is apparent in Fig. 2, where we plot the state of the system (blue) alongside the state estimate (dashed red) with error bounds (red shading). The bounds naturally accommodate the extreme noise levels, but immediately tighten when no measurement errors are present. Moreover, the map $\mathcal{L}$ from Assumption 1 naturally incorporates the system dynamics, which is why at some particular times the bounds increase, despite no measurement errors being added.

B. Ackerman Steering Model

We also consider a more physical system for a two-axle Ackerman steering model with "GPS" position outputs,

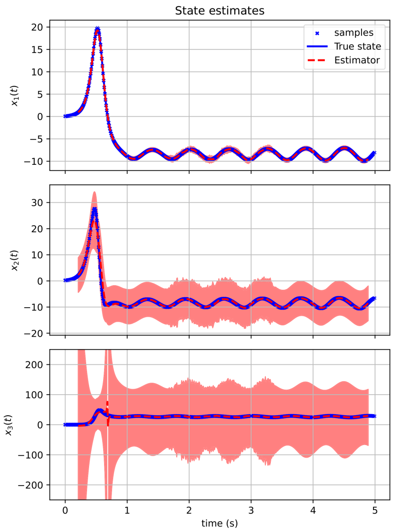

Fig. 2. State estimates for the Lorenz system. The true state is shown in blue, with dashed red lines and red shading indicating the state estimate and online error bounds of Corollary 1. Note that the system produces a *singular measurement* around $t = 0.5$.

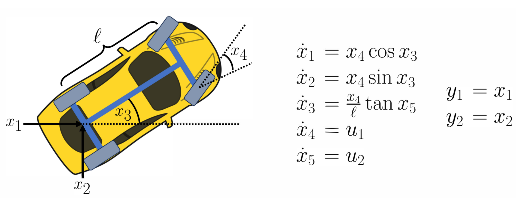

Fig. 3. A diagram illustrating the states of the Ackerman steering model.

whose states and dynamics are illustrated in Fig. 3.

We set the axle separation ℓ to 0.5, and use a sampling frequency of 100Hz ($\delta = 0.01$), fitting a degree $d = 5$ polynomial to the data using a sliding window of 50 measurements. We inject bounded measurement errors with magnitude $E = 0.025$ only in the interval $t \in [4.9, 9.8]$.

Here we show the state estimates (with error bounds) in Fig. 4. Notably, in the interval where there were measurement errors, the bounds inflate slightly and become less "smooth". The local dynamics of the vehicle, however, impact how much inflation occurs.

V. Conclusions

In this paper, we presented new deterministic worst-case error guarantees for a nonlinear state estimation scheme. Most importantly, our error bounds are easy to compute online and shrink or grow depending on the system behavior. These new error bounds directly interface with existing measurement-robust control frameworks, reducing the conservative nature of these methods.

We validated this estimator and its guarantees with two different nonlinear systems, verifying their performance and tightness. In the future, this principle of relating fit "residuals" to estimation errors could be extended to more classical estimators, proving new "online" error bounds for other families of nonlinear observers.

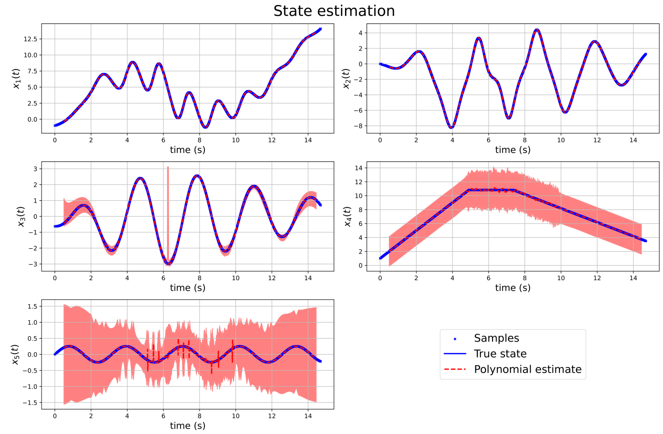

Fig. 4. State estimates for the Ackerman model. The true state is shown in blue, with dashed red lines and red shading indicating the state estimate and online error bounds of Corollary 1. The spike in x_3 at $t \approx 6$ is caused by angle wrapping artifacts. Spikes in x_5 *are* caused by singular measurements as estimating x_5 computes the *curvature* of the vehicle path.

References

[1] H. Khalil, *Nonlinear Control*, 1st ed. Pearson Education, 2015.

[2] P. Freeman, R. Kokotović, *Robust Nonlinear Control Design: State Space and Lyapunov Techniques*, 1st ed., ser. Modern Birkhäuser Classics. Boston, MA: Birkhäuser, 1996.

[3] R. Cosner, A. Singletary, A. Taylor, T. Molnar, K. Bouman, and A. Ames, "Measurement-robust control barrier functions: Certainty in safety with uncertainty in state," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2021, pp. 6286–6291.

[4] D. Agrawal and D. Panagou, "Safe and robust observer-controller synthesis using control barrier functions," *IEEE Control Systems Letters*, vol. 7, pp. 127–132, 2023.

[5] H. Shim and D. Liberzon, "Nonlinear observers robust to measurement disturbances in an iss sense," *IEEE Transactions on Automatic Control*, vol. 61, no. 1, pp. 48–61, 2016.

[6] M. Khajenejad and S. Yong, "Hinf-optimal interval observer synthesis for uncertain nonlinear dynamical systems via mixed-monotone decompositions," *IEEE Control Systems Letters*, vol. 6, pp. 3008–3013, 2022.

[7] J. Shamma and K. Tu, "Set-valued observers and optimal disturbance rejection," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 253–264, 1999.

[8] P. Bernard, *Observer design for nonlinear systems*. Springer, 2019, vol. 479.

[9] S. Diop, V. Fromion, and J. W. Grizzle, "A resettable kalman filter based on numerical differentiation," in *2001 European Control Conference*, 2001, pp. 1239–1244.

[10] M. Krichman, E. D. Sontag, and Y. Wang, "Input-output-to-state stability," *SIAM Journal on Control and Optimization*, vol. 39, no. 6, pp. 1874–1928, 2001.

[11] R. Freeman, "Global internal stabilizability does not imply global external stabilizability for small sensor disturbances," *IEEE Transactions on Automatic Control*, vol. 40, no. 12, pp. 2119–2122, 1995.

[12] S. Dashkoviskiy and L. Naujok, "Quasi-ISS/ISDS observers for interconnected systems and applications," *Systems & Control Letters*, vol. 77, pp. 11–21, 2015.

[13] S. Diop, "Observers for sampled data nonlinear systems via numerical differentiation," in *2007 European Control Conference*, 2007, pp. 1179–1184.

[14] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

[15] J. Grizzle and P. Moraal, "Newton, observers and nonlinear discrete-time control," in *29th IEEE Conference on Decision and Control*, 1990, pp. 760–767 vol.2.

[16] J. Berrut and L. Trefethen, "Barycentric lagrange interpolation," *SIAM Review*, vol. 46, no. 3, pp. 501–517, 2004.

[17] J. Bunton. (2023, Aug) Polynomial least-squares. UCLA. [Online]. Available: https://github.com/buntonj/polynomial_observers/blob/main/Polynomial_Regression_Bounds.pdf