Interval Observers for Simultaneous State and Model Estimation of Partially Known Nonlinear Systems

Mohammad Khajenejad, Zeyuan Jin, Sze Zheng Yong

Abstract—We study the problem of designing interval-valued observers that simultaneously estimate the system state and learn an unknown dynamic model for partially unknown nonlinear systems with dynamic unknown inputs and bounded noise signals. Leveraging affine abstraction methods and the existence of nonlinear decomposition functions, as well as applying our previously developed data-driven function over-approximation/abstraction approach to over-estimate the unknown dynamic model, our proposed observer recursively computes the maximal and minimal elements of the estimate intervals that are proven to contain the true augmented states. Then, using observed output/measurement signals, the observer iteratively shrinks the intervals by eliminating estimates that are not compatible with the measurements. Finally, given new interval estimates, the observer updates the over-approximation of the unknown model dynamics. Moreover, we provide sufficient conditions for uniform boundedness of the sequence of estimate interval widths, i.e., stability of the designed observer, in the form of tractable (mixed-)integer programs with finite countable feasible sets.

I. INTRODUCTION

Motivation. Motivated by the need to ensure safe and smooth operation in many safety-critical engineering applications such as fault detection, urban transportation, attack (unknown input) mitigation and detection in cyber-physical systems and aircraft tracking [1]–[3], robust algorithms for state and input estimation have been recently applied to derive compatible estimates of states and unknown inputs. Particularly, set/interval membership approaches have been broadly used to guarantee hard accuracy bounds in safety-critical bounded-error settings. Further, in practical systems, the existence of potentially dynamic unknown inputs with unknown dynamics makes the entire setting a partially unknown system. Thus, the development of appropriate data-driven methods that can deal with the noisy estimated data obtained form set/interval membership approaches to estimate/approximate/abstract unknown system models is a critical and interesting problem.

Literature review. Multiple approaches have been proposed in the literature to design set/interval observers [3]–[19], including linear time-invariant (LTI) [10], linear parameter-varying (LPV) [12], [16], Metzler and/or partial linearizable [9], [11], cooperative [8], [9], Lipschitz nonlinear [13], monotone nonlinear [6], [7] and uncertain nonlinear [14] systems. However, the aforementioned works either do not consider unknown inputs (i.e., input, disturbance, attack or noise signals with unknown dynamics to be reconstructed/estimated) [4]–[15], or the (potentially unbounded) unknown inputs do not affect the output (measurement) equation [16]. Considering systems where both state and output equations are affected by arbitrary unknown inputs, the problem of simultaneously designing state and unknown input “set-valued” observers has been studied in our previous works for LTI [3], LPV [17], switched linear [18] and nonlinear [19] systems with bounded-norm noise, while in our recent work [20], we particularly designed “interval-valued” observers for Lipschitz mixed-monotone nonlinear systems affected by arbitrary unknown inputs.

On the other hand, considering set-valued uncertainties, data-driven approaches that use sampled/observed input-output data to abstract or over-approximate unknown dynamics using a bounded-error setting, have gained increased popularity over the last few years [21]–[25]. The general objective of such data-driven methods is to find a set of known systems that share the most properties of interest with the unknown system dynamics [21], [22], under the assumption that the unknown dynamics is univariate Lipschitz continuous [23], multivariate Lipschitz continuous [24] or Hölder continuous [25]. Nonetheless, to our knowledge, these approaches do not explicitly deal with noise/disturbance and their effect on the abstraction, which is especially critical when dealing with “estimated” data. Hence, in our previous work [26], we generalized the aforementioned data-driven approaches to develop an abstraction approach that can use the noisy sampled/observed/estimated data to over-approximate the unknown Lipschitz continuous dynamics with upper and lower functions.

Contributions. The goal of this paper is to bridge between model-based set/interval-valued observer design approaches, e.g., in [3]–[19] and data-driven function approximation methods, e.g., in [21]–[25], to design interval-valued observers for nonlinear dynamical systems with bounded noise and dynamic unknown inputs, where the state and observation vector fields belong to a fairly general class of nonlinear functions and the unknown input dynamics is governed by an unknown input function. By extending the observer design approach in [20], we include a crucial update step, where starting from the intervals from the propagation step, the framers are iteratively updated by computing their intersection with the augmented state intervals that are compatible with the observations, resulting in the decreased width updated framers, which leads to obtain tighter intervals.

M. Khajenejad and S.Z. Yong are with the School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA (e-mail: {mkhajene, szyong}@asu.edu).

This work is partially supported by NSF grant CNS-1932066.
Moreover, by assuming a mild assumption of Lipschitz continuity for the unknown input functions and applying our previous data-driven function approximation/abstraction approach [26] to recursively over-approximate the unknown input function from the noisy estimated intervals/data obtained from the update step, as well as leveraging the combination of nonlinear decomposition/bounding functions [20], [27]–[29] and affine abstractions [30], we prove that our observer is correct, i.e., the framer property [11] holds and our estimation/abstraction of the unknown input model becomes more precise/tighter over time. More importantly, we provide sufficient conditions, in the form of tractable (mixed-)integer programs with finitely countable feasible sets, for the stability of our observer (i.e., the uniform boundedness of the sequence of estimate intervals widths). Further we compute uniformly bounded and convergent upper intervals for the sequence of estimates and derive their steady-state values.

II. PRELIMINARIES

Notation. \( \mathbb{R}^n \) denotes the \( n \)-dimensional Euclidean space and \( \mathbb{R}_{++} \) positive real numbers. For vectors \( v, w \in \mathbb{R}^n \) and a matrix \( M \in \mathbb{R}^{n \times n} \), \( \| v \| \triangleq \sqrt{v^T v} \) and \( \| M \| \) denote their (induced) 2-norm, and \( v \preceq w \) is an element-wise inequality. Moreover, the transpose, Moore-Penrose pseudoinverse, \((i,j)\)-th element and rank of \( M \) are given by \( M^T, M^\dagger, M_{i,j} \) and \( \text{rank}(M) \), while \( M_{(i,j)} \) is a sub-matrix of \( M \), consisting of its \( r \)-th through \( s \)-th rows. We call \( M \) a non-negative matrix, i.e., \( M \geq 0 \), if \( M_{i,j} \geq 0, \forall i \in \{1, \ldots, p\}, \forall j \in \{1, \ldots, q\} \). Also, \( M^+, M^{++} \in \mathbb{R}^{n \times n} \) are defined as \( M^+_{i,j} = M_{i,j} \) if \( M_{i,j} \geq 0 \), \( M^+_{i,j} = 0 \) if \( M_{i,j} < 0 \), \( M^{++} = M^+ + M^+ \). Furthermore, \( r = \text{rowsupp}(M) \in \mathbb{R}^p \), where \( r_i = 0 \) if the \( i \)-th row of \( A \) is zero and \( r_i = 1 \) otherwise, \( \forall i \in \{1, \ldots, p\} \). For a symmetric matrix \( S \), \( S \succeq 0 \) and \( S \preceq 0 \) \((S \succeq 0 \) and \( S \preceq 0 \) are positive and negative (semi-)definite, respectively. Next, we introduce some definitions and related results that will be useful throughout the paper.

Definition 1 (Interval, Maximal and Minimal Elements, Interval Width). An (multi-dimensional) interval \( I \subseteq \mathbb{R}^n \) is the set of all real vectors \( x \in \mathbb{R}^n \) that satisfies \( \underline{x} \leq x \leq \overline{x} \), where \( \underline{x}, \overline{x} \) and \( |\overline{x} - \underline{x}| \) are called minimal vector, maximal vector and width of \( I \), respectively.

Proposition 1. [13, Lemma 1] Let \( A \in \mathbb{R}^{n \times n} \) and \( x \preceq x \leq \overline{x} \in \mathbb{R}^n \). Then, \( A^+ \underline{x} - A^+ \overline{x} \preceq Ax \preceq A^+ \underline{x} - A^+ \overline{x} \). As a corollary, if \( A \) is non-negative, then \( Ax \preceq A \underline{x} \).

Definition 2 (Lipschitz Continuity). A vector field \( q(\cdot) : \mathbb{R}^n \to \mathbb{R}^m \) is \( L_q \)-Lipschitz continuous on \( \mathbb{R}^n \), if \( \exists L_q \in \mathbb{R}_{++}, \| q(\zeta_1) - q(\zeta_2) \| \leq L_q \| \zeta_1 - \zeta_2 \|, \forall \zeta_1, \zeta_2 \in \mathbb{R}^n \).

Definition 3 (Mixed-Monotone Mappings and Decomposition Functions). [27, Definition 4] A mapping \( f : \mathcal{X} \subseteq \mathbb{R}^n \to \mathcal{T} \subseteq \mathbb{R}^m \) is mixed monotone if there exists a decomposition function \( f_d : \mathcal{X} \times \mathcal{X} \to \mathcal{T} \) satisfying:

1. \( f_d(x, x) = f(x) \),
2. \( x_1 \geq x_2 \Rightarrow f_d(x_1, y) \geq f_d(x_2, y), \) and
3. \( y_1 \geq y_2 \Rightarrow f_d(x, y_1) \leq f_d(x, y_2) \).

Proposition 2. [28, Theorem 1] Let \( f : \mathcal{X} \subseteq \mathbb{R}^n \to \mathcal{T} \subseteq \mathbb{R}^m \) be a mixed monotone mapping with decomposition function \( f_d : \mathcal{X} \times \mathcal{X} \to \mathcal{T} \) and \( \underline{x} \preceq x \preceq \overline{x}, \) where \( \underline{x}, \overline{x} \in \mathcal{X} \). Then \( f_d(\underline{x}, \overline{x}) \leq f_d(\overline{x}, \underline{x}) \).

Note that the decomposition function of a vector field is not unique and a specific one is given in [27, Theorem 2].

System Assumptions. Consider the nonlinear discrete-time system with unknown inputs and bounded noise

\[ x_{k+1} = f(x_k, d_k, u_k, w_k), \]
\[ y_k = g(x_k, d_k, u_k, v_k), \]

where \( x_k \in \mathcal{X} \subseteq \mathbb{R}^n \) is the state vector at time \( k \in \mathbb{N} \), \( u_k \in U \subseteq \mathbb{R}^m \) is a known input vector, \( d_k \in D \subseteq \mathbb{R}^p \) is an unknown dynamic input vector that its dynamics is governed by an unknown vector field \( h(\cdot) \) as

\[ d_{k+1} = h(x_k, d_k, u_k, w_k), \]

and \( y_k \in \mathcal{Y} \subseteq \mathbb{R}^l \) is the measurement vector. The process noise \( w_k \in \mathbb{R}^n \) and the measurement noise \( v_k \in \mathbb{R}^l \) are assumed to be bounded, with \( w_k \leq \overline{w} \leq w \) and \( v_k \leq \overline{v} \leq v \), where \( \overline{w}, \overline{v} \) and \( \underline{w}, \underline{v} \) are the known lower and upper bounds of the process and measurement noise signals, respectively.

We also assume that lower and upper bounds, \( \underline{d}_0 \) and \( \overline{d}_0 \), for the initial augmented state \( z_0 \triangleq [x_0^T d_0^T] \) are available, i.e., \( \underline{z}_0 \leq z_0 \leq \overline{z}_0 \). The vector fields \( f(\cdot) : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}^n \) and \( g(\cdot) : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}^l \) are known, while the vector field \( h(\cdot) : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}^p \) is unknown, but each of its arguments \( h_j(\cdot) : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \to \mathbb{R}, j \in \{1, \ldots, p\} \) is known to be Lipschitz continuous with the known Lipschitz constant \( L^h_j \). Moreover, we assume the following:

Assumption 1. Vector field \( f(\cdot) \) is mixed-monotone with decomposition function \( f_d(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n \).
Assumption 2. The entire space \( X = \mathbb{Z} \times U \) is bounded, where \( \mathbb{Z} = X \times D \) and \( U \) are the spaces of the augmented states \( z_k \triangleq [x_k^T \, d_k^T]^T \) and the known inputs \( u_k \), \( \forall k \in \{0,\ldots,\infty\} \), respectively.

Note that Assumption 1 is satisfied for a broad range of nonlinear functions [29], while Assumption 2 is reasonable for most practical systems.

The observer design problem can be stated as follows:

Problem 1. Given a partially known nonlinear discrete-time system (3) with bounded noise signals and unknown dynamic inputs (5), design a stable observer that simultaneously finds bounds of interval-valued observers to the system for most practical systems.

IV. STATE AND MODEL INTERVAL OBSERVERS (SMIO)

A. Interval-Valued Recursive Observer

A three-step recursive interval-valued observer that combines model-based and data-driven approaches will be considered in this paper. The observer structure is composed of a State Propagation (SP), a Measurement Update (MU) step and a Model Learning (ML) step. In the state propagation step, the interval for the augmented states (consisting of the state and the unknown input) is propagated for one time step through the nonlinear state equation and the upper and lower approximation of the unknown input function obtained in previous time step. In the update step, compatible intervals of the augmented states are iteratively updated given new measurements and observation function, and finally the model learning step re-estimates the upper and lower approximations (abstractions) for the function of the unknown inputs. More formally, the three observer steps have the following form (with \( z_k \triangleq [x_k^T \, d_k^T]^T \), \( z_k^p \triangleq [x_k^p \, d_k^p]^T \), \( z_k^q \triangleq [x_k^q \, d_k^q]^T \), etc.):

**SP:** \( \tilde{T}_k^p = F^p(\tilde{T}_{k-1}^p, y_{k-1}, u_{k-1}, e_{k-1}), \tilde{h}_{k-1}(-), \tilde{b}_{k-1}(-) \),

**MU:** \( \tilde{T}_k^p = F^p(\tilde{T}_k^p, y_k, u_k) \),

**ML:** \( \tilde{h}_k(-), \tilde{b}_k(-) \) being to-be-designed interval-valued mappings and \( F^p \) a to-be-constructed function over-approximation process (abstraction model), while \( \tilde{T}_k^p \) and \( \tilde{T}_k^q \) are the intervals of compatible estimated and augmented states and \( \{\tilde{h}_k(-), \tilde{b}_k(-)\} \) is a data-driven abstraction over-approximation model for the unknown function \( h(-) \), at time step \( k \), respectively, i.e., \( \forall \tilde{z}_k \in D_h : \tilde{h}_k(\tilde{z}_k) \leq h(\tilde{z}_k) \leq \tilde{b}_k(\tilde{z}_k) \) at time step \( k \), where \( D_h \) is the domain of \( h(-) \) and \( \tilde{z}_k \triangleq [x_k^T \, u_k^T]^T \).

To leverage the properties of intervals [16], while taking into consideration the computational complexity of optimal observers [31], we consider the following form of interval estimates in the propagation and update steps:

\[ \tilde{T}_k^p = \{z \in \mathbb{R}^{n+p} : \tilde{z}_k^p \leq z \leq \tilde{z}_k^q\}, \]

\[ \tilde{z}_k = \{z \in \mathbb{R}^{n+p} : \tilde{z}_k \leq \tilde{z}_k^q \leq \tilde{z}_k \}, \]

where the estimation boils down to find the maximal and minimal values of \( \tilde{T}_k^p \) and \( \tilde{T}_k^q \), i.e., \( \tilde{T}_k^p, \tilde{T}_k^q, \tilde{z}_k, \tilde{z}_k \). Further, at the model learning step, given the interval estimates for a certain period of time as data, we use a data-driven function abstraction/over-approximation model, developed in our previous work [26], to update our previously estimated model of the input dynamics \( h(-) \) in the current time step.

B. Observer's Structure

Our interval observer can be defined at each step \( k \geq 1 \) as follows (with augmented state \( z_k \triangleq [x_k^p \, d_k^p]^T \), \( \zeta_k \triangleq [z_k^p \, u_k \, w_k]^T \) and known \( \tilde{a}_0 \) and \( \tilde{x}_0 \) such that \( \tilde{a}_0 \leq x_0 \leq \tilde{a}_0 \)):

**State Propagation (SP):**

\[ \begin{align*}
\tilde{T}_k^p &= \min \left( f_d(\tilde{T}_{k-1}^p, u_{k-1}, \tilde{w}_{k-1}, \tilde{z}_{k-1}, \tilde{w}_{k-1}) \right) \\
&= \max \left( f_d(\tilde{T}_{k-1}^q, u_{k-1}, \tilde{w}_{k-1}, \tilde{z}_{k-1}, \tilde{w}_{k-1}) \right)
\end{align*} \]

**Measurement Update (MU):**

\[ \tilde{z}_k = \left[ \begin{array}{c} \tilde{z}_{k,1}^p \\ \tilde{z}_{k,1}^q \end{array} \right] = \left[ \begin{array}{c} \tilde{z}_{k,1}(1,n) \\ \tilde{z}_{k,1}(n+1:n+p) \end{array} \right], \]

**Model Learning (ML):**

\[ \begin{align*}
\tilde{h}_{k,j}(\zeta_k) &= \min_{t \in \{1,\ldots,n\}} |\tilde{T}_{k-1}^q - L_j^P h(\tilde{z}_{k-1})| + \varepsilon_{k-1}^+ \\
&= \max_{t \in \{1,\ldots,n\}} |\tilde{T}_{k-1}^q - L_j^P h(\tilde{z}_{k-1})| + \varepsilon_{k-1}^-
\end{align*} \]

where \( j \in \{1,\ldots,p\} \), \( \{\tilde{T}_{k-1}^q - L_j^P h(\tilde{z}_{k-1})\}_{i=0} \) are the augmented input-output data set.

At each time step \( k \), the augmented data set constructed from the estimated framers gathered from the initial to the current time step, is used in the model learning step to recursively derive over-approximations of the unknown function \( h(-) \), i.e., \( \{\tilde{h}_k(-), \tilde{b}_k(-)\} \) by applying [26, Theorem 1]. In addition

\[ \tilde{T}_k^{p} = \tilde{T}_k^{p} + \tilde{T}_k^{q} = \tilde{T}_k^{p} + \tilde{T}_k^{q} + \tilde{T}_k^{p} + \tilde{T}_k^{q} \]

**Model Learning (ML):**

\[ \begin{align*}
\tilde{T}_k^{p} &= \tilde{T}_k^{p} + \tilde{T}_k^{q} + \tilde{T}_k^{p} + \tilde{T}_k^{q} \\
&= \tilde{T}_k^{p} + \tilde{T}_k^{q} + \tilde{T}_k^{p} + \tilde{T}_k^{q}
\end{align*} \]

Furthermore, \( \forall q \in \{f, h\}, \forall i \in \{A, W\}, \zeta_k \in \{1,\ldots,\infty\}, j \in \{1,\ldots,p\}, \omega_{i,k} \in A^P_q, B^P_q, W^P_q, \tilde{a}_0, \tilde{e}_0 \), \( \tilde{e}_0 \), \( \varepsilon_{k-1}^+ \) and \( f_d(\ldots) \) are to-be-designed observer parameters.
Algorithm 1 SMIO

1: Initialize: maximal($\overline{z}_0$) = $\overline{z}_0$; minimal($\underline{z}_0$) = $\underline{z}_0$;
    $\triangleright$ Observer Gains Computation
  \forall q_i \in \{f, h\}, \exists j \in \{A, B, W\}, i \in \{1, \ldots, 2p\}, j \in \{1, \ldots, p\}
  compute $\omega_{ik}, A^q_{ik}, B^q_{ik}, W^q_{ik}$, $\overline{z}_{ik}^q$, $\underline{z}_{ik}^q$, $\overline{z}_{ik}^q$, $\underline{z}_{ik}^q$ via Theorem 1 and Appendix VI-A.
2: for $k = 1$ to $\infty$ do
    $\triangleright$ Augmented State Estimation
    Compute $\overline{z}_{ik}^q$ via (4a)–(4e) and \{\overline{z}_{ik}^q, \underline{z}_{ik}^q\}$_{0}^{\infty}$ via (8)–(11);
    (3) $\overline{z}_{ik} = \overline{z}_{ik}^q, \underline{z}_{ik} = \underline{z}_{ik}^q$; $\overline{z}_k = \{z \in \mathbb{R}^n \mid \overline{z}_k \leq z \leq \underline{z}_k\}$;
    Compute $\delta_k^q$ through Lemma 3
    $\triangleright$ Model Estimation
    Compute $\check{h}_k(\cdot), \underline{h}_k(\cdot)$ via (5a)–(5b);
4: end for

Note that since the tightness of the upper and lower bounding functions for the observation function $q$ (cf. Propositions 3 and 2) depends on the \textit{a priori} interval $B$, the measurement update step is done iteratively (see proof of Theorem 2 for more explanation). Hence, if tighter updated intervals are obtained starting from the compatible intervals from the propagation step, we can use them as the new $B$ to obtain better abstraction/bounding functions for $q$, which in turn may lead to even tighter updated intervals. Repeating this process results in a sequence of monotonically tighter updated intervals, that is convergent by the monotone convergence theorem, and its limit is chosen as the final interval estimate at time $k$.

Further, benefiting from our previous result in [26, Theorem 1], where we developed a data-driven approach for over-approximation/abstraction of Lipschitz unknown nonlinear functions given noisy data, in the model learning step, we treat the history of obtained compatible intervals in the past time steps up to the current time, $\{[\underline{z}_k, \overline{z}_k]\}$ as the noisy input data and the compatible interval of unknown inputs, $[\overline{d}_k, \underline{d}_k]$, as the noisy output data to recursively construct a sequence of abstraction/over-approximation models $\{\check{h}_k(\cdot), \underline{h}_k(\cdot)\}$ for the unknown input function $h(\cdot)$, that by construction satisfy (13), i.e., our input model estimation is correct and becomes more precise over time (cf. Lemma 1). Algorithm 1 summarizes the SMIO observer.

C. Correctness of the Observer

The objective of this section is to design the SMIO observer’s gains such that the \textit{framer property} [11] holds, i.e., we desire to guarantee that the observer returns correct interval estimates, in the sense that starting from the initial interval $\overline{z}_0 \leq z_0 \leq \underline{z}_0$, the true augmented states of the dynamic system (2) are guaranteed to be within the estimated intervals, given by (4a)–(5b). If the observer is correct, we call $\{\overline{z}_k, \underline{z}_k\}$ an augmented state framer sequence for system (2).

Before deriving our main first result on correctness of the observer, we state a modified version of our previous result in [30, Theorem 1], in a unified manner that enables us to derive parallel global and local affine bounding functions for our known $f(\cdot), g(\cdot)$ and unknown $h(\cdot)$ vector fields.

Proposition 3 (Parallel Affine Abstractions). Let the entire space be defined as $\mathbb{X}$ and suppose that Assumption 2 holds.

Consider the vector fields $\overline{q}(\cdot), q(\cdot) : \mathbb{X} \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, where $\forall \zeta \in \mathbb{X}, \overline{q}(\zeta) \leq q(\zeta)$, along with the following Linear Program (LP):
\[\begin{align*}
\min_{\theta^q_B, A^q_B, Z_{Bq}} & \quad \theta^q_B \\
\text{s.t.} & \quad A^q_B \overline{q}(z) + \varepsilon^q_B \leq q(z) \leq A^q_B \overline{q}(z) + \overline{q}(z), \\
& \quad \overline{q}(z) - \overline{q}(z) \leq (A^q_B - A^q_B) \zeta \leq \overline{q}(z) - \overline{q}(z),
\end{align*}\]
where $B$ is an interval with $\overline{X}, \overline{X}$ being its maximal, minimal and set of vertices, respectively, $1_m \in \mathbb{R}^m$ is a vector of ones, $\sigma^q$ is given in [30, Proposition 1 and (8)] for different classes of vector fields and ($A^q, \overline{q}, \overline{q}, \varepsilon^q$) is the global parallel affine abstraction matrices for the pair of functions $\overline{q}(\cdot), q(\cdot)$ on the entire space $\mathbb{X}$, i.e.,
\[\begin{align*}
\overline{A}^q \zeta + \varepsilon^q \leq q(\zeta) \leq \overline{A}^q \zeta + \overline{q}(\zeta), \forall \zeta \in \mathbb{X}. \tag{13}
\end{align*}\]

Suppose that ($A^q, \overline{q}, \overline{q}, \varepsilon^q$) are not known. Then, solving (12a) on the entire space $\mathbb{X}$, i.e., when $B = \mathbb{X}$ (where the constraint (12b) is trivially satisfied and is thus redundant) returns a tuple of ($\theta^q, A^q, A^q, \varepsilon^q$) that satisfies (13). i.e., constructs a global affine abstraction model for the pair of functions $\overline{q}(\cdot), q(\cdot)$ on the entire space $\mathbb{X}$.

Now, suppose that ($A^q, \overline{q}, \overline{q}, \varepsilon^q$) are known (or have been computed as described above). Then, solving (12a) on $B$ constrained to (12b), returns a tuple of parallel affine abstraction matrices for the pair of functions $\overline{q}(\cdot), q(\cdot)$ on the interval $B$, satisfying the following: $\forall \zeta \in \mathbb{B}$
\[\begin{align*}
\overline{A}^q \zeta + \varepsilon^q \leq A^q_B \zeta + \varepsilon^q_B \leq q(z) \leq A^q_B \zeta + \overline{q}(z) \leq \overline{A}^q \zeta + \overline{q}(z), \tag{14}
\end{align*}\]

Next, equipped with all the required tools, we state our first main result on the framer property of the SMIO observer.

Theorem 1 (Correctness of the Observer). Consider the system (2) with its augmented state defined as $z = [x^T \; a^T]^T$, along with the SMIO observer in (4a)–(5b). Suppose that Assumptions 1–2 hold, $f_q(\cdot)$ is a decomposition function of $f(\cdot)$ and observer gains and parameters are designed as given in (VI-A). Then, the SMIO observer is correct, i.e., the sequences $\{\overline{z}_k, \underline{z}_k\}$ construct framers for the augmented state sequence of system (2).

Next, we show that given correct interval estimates, the abstraction model of the unknown input function becomes tighter (i.e., more precise) over time, so our estimation of the unknown input model becomes more accurate over time.

Lemma 1. Consider the system (2) and the SMIO observer in (4a)–(5b) and suppose that all the assumptions in Theorem 1 hold. Then, the following holds:
\[\begin{align*}
\underline{h}_0(z_0) \leq \ldots \leq \underline{h}_k(z_k) \leq \ldots \leq \lim_{k \rightarrow \infty} \underline{h}_k(z_k) \leq \underline{h}(z_k) \leq \underline{h}(z_k) \leq \lim_{k \rightarrow \infty} \overline{h}_k(z_k) \leq \ldots \leq \overline{h}_0(z_0), \tag{15}
\end{align*}\]
i.e., the unknown input model estimations/abstractions are correct and become more precise/tighter in time.

D. Observer Stability

In this section, we study the stability of the designed observer. We first formally define the notion of stability that
we investigate in this paper.

**Definition 4** (Stability). The observer SMIO **(4a)–(6b)** is stable, if the interval widths sequence \( \{||\Delta z_k||\} \) is uniformly bounded, and consequently the sequence of estimation errors \( \{||z_k-z_k^-||\} \) is also uniformly bounded.

Next, we derive a property for the decomposition function given in **(1)**, which will be helpful in deriving sufficient conditions for the observer’s stability.

**Lemma 2.** Let \( q(\cdot): \mathbb{X} \subset \mathbb{R}^n \to \mathbb{R}^m \) be a mixed-monotone vector-field with a corresponding decomposition function \( q_d(\cdot) \) constructed using **(1)**. Suppose that Assumption **2** holds and let \( (A^q, \tau^q, \epsilon^q) \) be the parallel affine abstraction matrices for function \( q(\cdot) \) on its entire domain \( \mathbb{X} \) (can be computed via Proposition **3**). Consider any ordered pair \( \zeta \leq \zeta' \in \mathbb{X} \). Then, \( \Delta q \leq (|A^q|+2C^q)|\Delta \zeta + \Delta \epsilon^q| \), where \( \Delta q \equiv q_d(\zeta, \zeta')-q_d(\zeta', \zeta) \), \( \| \Delta \zeta \| \leq \| \zeta' \zeta \| \) and \( C^q \) given in **(1)**.

Now we are ready to state our next main result on the SMIO observer’s stability through the following theorem.

**Theorem 2** (Observer Stability). Consider the system **(2)** along with the SMIO observer **(4a)–(6b)**. Let \( D_n \) be the set of all diagonal matrices in \( \mathbb{R}^{n \times n} \) with their diagonal arguments being 0 or 1. Suppose that all the assumptions in Theorem **2** hold and the decomposition function \( f_d \) is constructed using **(1)**. Then, the observer is stable if

\[
L^+ \triangleq \min_{(D_1, D_2, D_3) \in D_n^3} ||A^q(D_1, D_2)A^{f,h}(D_3)|| \leq 1, \quad (16a)
\]

subject to

\[
D_{1,i,i} = 0, \quad i = 1, \quad (16b)
\]

with

\[
A^q(D_1, D_2) \triangleq (I - D_1) + D_1A^q|I - D_2|A^q, \\
A^{f,h}(D_3) \triangleq [[A^f + 2(I - D_3)C_f^r]A^h|A^h] + \{A^h \} \triangleq \bigcup_{k=0}^{\infty} a^h_{k+1}(I_n + p^h_k), \quad k \quad \text{given in Proposition 3}, \quad r \equiv \text{routsupp}(I - A^q), C_f^r \triangleq C_f^rC_f^r, \quad \text{given in (1)}
\]

and \( D_n^+ \triangleq D_n \cup D_n^0 \). Further,

\[
f_1(\zeta_k) = 0.6x_{1,k} - 0.12x_{2,k} - 0.2x_{1,k}, \quad f_2(\zeta_k) = -0.2x_{1,k} - 0.14x_{2,k}, \quad d_{1,k+1} = 0.1\cos(d_{1,k}), \quad g_{1}(v_k) = 0.2x_{1,k} + 0.65x_{2,k} + 0.8\sin(0.3x_{1,k} + 2x_{2,k}), \quad g_2(v_k) = \sin(x_{1,k}, d_{2,k+1} = 1 - x_{1,k}^2 - 0.1d_{1,k},
\]

with \( u_k^T \triangleq [u_k^T u_k^T v_k^T] \). Moreover, using Proposition **3** while abstraction slopes are set to zero, we can obtain finite-valued upper and lower bounds (horizontal abstractions) for the partial derivatives of \( f(\cdot) \) as:

\[
|a_{11}^{f_1} a_{12}^{f_1} a_{21}^{f_2} a_{22}^{f_2}| = \begin{bmatrix}
0.38 & -0.46 \\
-0.2 - \epsilon & -0.14 - \epsilon
\end{bmatrix}, \quad |b_{11}^{f_1} b_{12}^{f_1} b_{21}^{f_2} b_{22}^{f_2}| = \begin{bmatrix}
0.82 & 0.21 \\
-0.2 + \epsilon & -0.14 + \epsilon
\end{bmatrix},
\]

where \( \epsilon \) is a very small positive value, ensuring that the partial derivatives are in open intervals (cf. [27, Theorem 1]). Therefore, Assumption **4** holds by [27, Theorem 1]. Hence, we expect that the true states and unknown inputs are within the estimate intervals by Theorem **1** i.e., the interval estimates are correct. This can be observed from Figure **1** where the true states and unknown inputs as well as interval estimates are depicted.

Furthermore, solving Proposition **3** for global abstraction matrices, we derive \( A^f = \begin{bmatrix}
0.4063 & 0.1706 & 0 & -0.1 & 0 \\
-0.2 & -0.14 & -0.2 & -0.2 & 0 & 1
\end{bmatrix} \), \( A^g = \begin{bmatrix}
0.4204 & 0.797 & -0.1 & 0.3 & 1 & 0 \\
0.584 & 0 & 0.5 & -0.7 & 0 & 1
\end{bmatrix} \), and from [27, (10)–(13)], we
versus global abstractions, with the cost of more extensive
is the main benefit of using iterative local affine abstractio ns
abstraction-based intervals are still uniformly bounded. This
results in
starting to implement the observer) for observer stability , we
turn, lead to updated local intervals that by construction a re
consequently, obtain the updated local abstractions, whic h, in
for the implementation, we iteratively update the framers a nd

Note that as discussed in the proof of Theorem 2 since
we need to check an a priori condition (i.e., offline or before
starting to implement the observer) for observer stability, we
use global abstraction slopes for stability analysis. However,
for the implementation, we iteratively update the framers and
consequently, obtain the updated local abstractions, which, in
turn, lead to updated local intervals that by construction are
tighter than the global ones, as shown in the proof of Theo-
rem 2 Hence, for a given system, it might be the case that the
(relatively conservative) global-abstraction-based sufficient
conditions for the observer stability given in Theorem 2 do
not hold, i.e., \( L^* > 1 \), while the implemented local-
abstraction-based intervals are still uniformly bounded. This
is the main benefit of using iterative local affine abstractions
versus global abstractions, with the cost of more extensive
computational effort. Figure 3 compares the tightness of

obtain \( C_f = \begin{bmatrix} 0.374 & 0.02 \\ 0.0135 & 0.407 \end{bmatrix} \), using (1). Consequently, the
mixed-integer program (16a) constrained by (16b)
results in \( L^* = 1.1 > 1 \) and so the sufficient conditions in
Theorem 2 are not satisfied. Despit this, as can be seen in
Figure 2 we obtain uniformly bounded and convergent
interval estimate errors when applying our observer design
procedure, where at each time step, the actual error sequence
is upper bounded by the interval widths, which converge to
steady-state values.

VI. CONCLUSION
An interval-valued observer for partially unknown non-
linear systems with dynamic unknown inputs and bounded
noise signals was designed in this paper, that simultaneously
estimated the augmented states and unknown input (with
unknown dynamics) of the system. By applying a combi-
nation of nonlinear bounding/decomposition functions and
affine abstractions as well as benefiting from our previously
developed data-driven function abstraction method to over-
estimate the unknown input model from the noisy estimated
input-output data, we showed that the estimate interval
estimates are correct in the sense that our proposed observer
dependently computes the maximal and minimal elements of
the estimate intervals that are proven to contain the true
augmented states, and by observing new output/measurement
signals, iteratively shrinks the intervals by eliminating es-
timates that are not compatible with the measurements.
Moreover, sufficient conditions for uniform boundedness of
the sequence of estimate interval widths, i.e., stability of the
designed observer were provided in the form of tractable
(mixed-)integer programs with finitely countable feasible
sets.

REFERENCES
[1] W. Liu and I. Hwang. Robust estimation and fault detection and
isolation algorithms for stochastic linear hybrid systems with unknown
fault input. IET control theory & applications, 5(12):1353–1368, 2011.
[2] S.Z. Yong, M. Zhu, and E. Frazzoli. Switching and data inj ection
attacks on stochastic cyber-physical systems: Modeling, resilient esti-
mation and attack mitigation. ACM Transactions on Cyber-Physical
Systems, 2(2):9, 2018.
[3] S.Z. Yong. Simultaneous input and state set-valued observers with
applications to attack-resilient estimation. In 2018 Annual American
Control Conference (ACC), pages 5167–5174. IEEE, 2018.
[4] L. Jaulin. Nonlinear bounded-error state estimation of continuous-time
systems. Automatica, 38(6):1079–1082, 2002.
[5] M. Kieffer and E. Walter. Guaranteed nonlinear state estimator for
coopetative systems. Numerical algorithms, 37(1-4):187–198, 2004.
[6] M. Moisan, O. Bernard, and J-L. Gouzé. Near optimal interval
observers bundle for uncertain bioreactors. In European Control
Conference (ECC), pages 5115–5122. IEEE, 2007.
[7] O. Bernard and J-L. Gouzé. Closed loop observers bundle for uncertain
biotechnological models. Journal of Process Control, 14(7):765–774, 2004.
APPENDIX: OBSERVER GAIN DEFINITIONS AND PROOFS

A. Observer Gain Definitions

\[ \forall j \in \mathcal{J}, j \in \{ f, h \}, j \in \{ A, W \}, i \in \{ 1, \ldots, \infty \} : \]
\[ J^q_k = \left[ J^T_k, J^F_k \right], B^q_k = B^T_k, B^F_k \right], E^q_k = \left[ E^T_k \right], F^q_k = F^T_k, F^F_k \]}
\[ \omega_{i,k} = \text{krwassup}(\mathcal{I} - A^{q}_k A^{f}_k) \]
\[ \left( A^{q}_k, B^{q}_k, W^{q}_k, \pi^{q}_k, \lambda^{q}_k \right), \left( A^{f}_k, B^{f}_k, W^{f}_k, \pi^{f}_k, \lambda^{f}_k \right) \] are solutions to the problem (12a) for the corresponding functions \( g(x) = \lambda(x) \) for \( f, \lambda \) and \( \mathcal{I}^{\lambda} \) and \( \lambda(x) \), respectively, at time \( k \) and iteration \( i \), while \( \alpha \) is a very large positive real number (infinity).

B. Proof of Proposition 5

Consider the case when the global affine abstraction matrices are unknown. Then, by setting \( \mathcal{B} = X \), \( A^{q}_k = h^k \) and \( \theta^{q}_k \), constraint (12b) is redundant and so the LP (12a) boils down to a special case of the LP in [30, 16], with only one considered partition. Then, (12) follows from [30, Theorem 1]. Moreover, given the global affine abstractions, solving the LP in (12a) is equivalent to solving the the LP in [30, 16] on the corresponding interval (set) of \( \mathcal{B} \), with the extra (non-trivial) constraint (12b). This constraint along with the result in [30, Theorem 1] result in (14).

C. Proof of Theorem 7

To use induction and as for the induction base, by assumption, \( z_0 \leq z_0 \leq z_{\alpha} \) holds. Now for the induction step, suppose that \( z_{k-1} \leq z_{k-1} \leq z_{k-1} \). Then, Propositions 1, 5 as well as (2) (7) (43) and [26, Theorem 1] imply that \( z_{k} \leq z_{k} \leq z_{k} \). Given this, iteratively obtaining upper and lower abstraction matrices for the observation function \( g(y) \) based on Proposition 3 and applying Proposition 1 result in
\[ \omega_{i,k} \leq A^{q}_k z_{k} \leq \omega_{i,k} \]
for \( \alpha_{i,k} \) are given in (11) and \( A^{q}_k \) is a solution of the LP in (12a) i.e., is a parallel abstraction slope for function \( g(y) \) at iteration \( i \) in the corresponding compatible interval \( [z_{k-1}^{\alpha} - z_{k-1}^{\alpha}] \). Then, multiplying (19) by \( A^{q}_k \), Proposition 1 the fact that \( z_{k-1}^{\alpha} - z_{k-1}^{\alpha} \) already construct framers for the augmented state \( z_{k} \) at time \( k \) and [33] imply that \( z_{k} \leq z_{k} \leq z_{k} \) with \( z_{k}^{\alpha} \) given in (9). Now, note that by construction, the sequences of updated upper and lower framers, \( \pi_{i,k}^{\infty} \) and \( \pi_{i,k}^{\infty} \) with \( \pi_{i,k}^{\infty} = \pi_{i,k}^{\infty} \) and \( \pi_{i,k}^{\infty} = \pi_{i,k}^{\infty} \) are monotonically decreasing and increasing, respectively, and hence are convergent by the monotone convergence theorem. Consequently, their limits \( \pi_{i,k} \). are the returned updated augmented state framers by the observer. This completes the proof.
D. Proof of Lemma 1

It directly follows from [26, Theorem 1] and Theorem 1 that the model estimates are correct, i.e., \( \forall k \in \{0, \ldots, \infty\} : h_k(\zeta_k) \leq h_k(\zeta_k) \leq h_k(\zeta_k) \). Moreover, considering the data-driven abstraction procedure in model learning step, note that by construction the data set used at time step \( k \) is a subset of the one used at time \( k+1 \). Hence, by [26, Proposition 2] the abstraction model satisfies monotonicity, i.e., (15) holds.

E. Proof of Lemma 2

Starting form (11), it is not hard to verify that
\[
\Delta q_{\zeta} = q(\zeta_1) - q(\zeta_2) + 2C^q \Delta \zeta, \tag{20}
\]
for some \( \zeta_1, \zeta_2 \) that satisfy \( \zeta \leq \zeta_1, \zeta_2 \leq \zeta \). On the other hand, by Proposition 3 in addition to Proposition 1 \( \forall j \in \{1, 2\} \):
\[
\Delta x^{\top} \zeta - \Delta x^{\top} \zeta + e^2 \leq q(j) \leq \Delta x^{\top} \zeta - \Delta x^{\top} \zeta + e^2,
\]
which implies \( q(\zeta_1) - q(\zeta_2) \leq |\Delta x| \Delta q_{\zeta} + |\Delta e| q_{\zeta}^{\top} \Delta \zeta. \) Combining this and (20) yields the result.

F. Proof of Theorem 2

Note that our goal is to obtain sufficient stability conditions that can be checked a priori instead of for each time step \( k \). On the other hand, for the implementation of the update step, we iteratively find new parallel affine abstraction slopes \( A^j_{k,k} \) by iteratively solving the LP (12a) for \( g \) on the intervals obtained in the previous iteration, \( B_{k,k} = [z^{u}_{k-1,k}, z^{l}_{k-1,k}] \), to find local framers \( \tilde{x}^{u}_{k,k}, \tilde{x}^{l}_{k,k} \) (cf. (8)–(11)), with additional constraints given in (23) in the optimization problems, which guarantees that the iteratively updated local intervals obtained using the local abstraction slopes are inside the global interval, i.e.,
\[
\tilde{x}^{u}_{k,k} \leq x^{u}_{k,k} \leq \cdots \leq x^{u}_{0,k} \leq \lambda_{k} \leq x^{l}_{0,k} \leq \cdots \leq x^{l}_{k,k} \leq \tilde{x}^{l}_{k,k},
\]
where \( x^{u}_{k,k}, \tilde{x}^{u}_{k,k} \) can be obtained by applying (9) for just one iteration (dropping index \( i \)) while \( \tilde{x}^{u}_{k,0} = \tilde{x}^{p}_{k,0} = 2k, \)
\[
\tilde{x}^{u}_{k,k} = \left[ \begin{array}{c}
\min(A^{g^{\top}+} \lambda_k + A^{g^{\top}+} \omega, \tilde{x}^{u}_{k,k})
\max(A^{g^{\top}+} \lambda_k + A^{g^{\top}+} \omega, \tilde{x}^{u}_{k,k})
\end{array} \right], \tag{21}
\]
This allows us to use the global parallel affine abstraction slope \( A^g \) for the stability analysis as follows. Dropping index \( i \) in (10)–(11) and defining \( \Delta_{\zeta}^{\top} = \zeta_k - \tilde{z}_k \) (and similarly for \( \Delta_{\tilde{z}}^{\top} \)), (9) implies that \( \forall D_1 \in D_{n+p} \):
\[
\Delta_{\zeta}^{\top} \leq \min(|A^{\theta^g}| \Delta_{\zeta}^{\top} + 2k \Delta t, \Delta_{\tilde{z}}^{\top}) \leq D_1(|A^{\theta^g}| \Delta_{\zeta}^{\top} + 2k \Delta t + (I - D_1) \Delta_{\zeta}^{\top}) \Delta_{\zeta}^{\top}, \tag{22}
\]
where the second inequality follows from generalization of the fact that \( \min(a, b) \leq \lambda a + (1 - \lambda) b, \forall a, b \in \mathbb{R}, \lambda \in [0, 1] \). Moreover, (10)–(11) and similar reasoning imply that \( \forall D_2 \in D_1 \):
\[
\Delta_{\zeta}^{\top} \leq \min(|W^g| \Delta v + \Delta_{\zeta}^{\top}, |A^g| \Delta_{\zeta}^{\top}) \leq D_2(|W^g| \Delta v + \Delta_{\zeta}^{\top}) + (I - D_2) |A^g| \Delta_{\zeta}^{\top}. \tag{23}
\]
On the other hand, by similar arguments, it follows from (14)–(15) that \( \forall D_3 \in D_n \):
\[
\Delta_{\tilde{z}}^{\top} \leq \left[ D_3(|A^f| \Delta_{\tilde{z}}^{\top} + |W| \Delta w + \Delta_{\tilde{z}}^{\top} + (I - D_3) \Delta_{\tilde{z}}^{\top}) \right]. \tag{24}
\]