

Safe Learning-based Observers for Unknown Nonlinear Systems using Bayesian Optimization

Chakrabarty, Ankush; Benosman, Mouhacine

TR2021-101 September 16, 2021

Abstract

Data generated from dynamical systems with unknown dynamics enable the learning of state observers that are: robust to modeling error, computationally tractable to design, and capable of operating with guaranteed performance. In this paper, a modular design methodology is formulated, that consists of three design phases: (i) an initial robust observer design that enables one to learn the dynamics without allowing the state estimation error to diverge (hence, safe); (ii) a learning phase wherein the unmodeled components are estimated using Bayesian optimization and Gaussian processes; and, (iii) a re-design phase that leverages the learned dynamics to improve convergence rate of the state estimation error. The potential of our proposed learning-based observer is demonstrated on a benchmark nonlinear system. Additionally, certificates of guaranteed estimation performance are provided.

Automatica 2021

© 2021 MERL. This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Safe Learning-based Observers for Unknown Nonlinear Systems using Bayesian Optimization

Ankush Chakrabarty^{a,*}, Mouhacine Benosman^{a,**}

^a*Mitsubishi Electric Research Laboratories, Cambridge, MA, USA.*

Abstract

In this paper, a modular observer design methodology is formulated for nonlinear systems with partial model knowledge. Our design consists of three design phases: (i) an initial robust observer design that enables one to learn the dynamics without allowing the state estimation error to diverge (hence, *safe*); (ii) a learning phase wherein the unmodeled components are estimated using Gaussian process-based Bayesian optimization; and, (iii) a re-design phase that leverages the learned dynamics to improve convergence rate of the state estimation error. The potential of our proposed learning-based observer is demonstrated on a benchmark nonlinear system. Additionally, certificates of guaranteed estimation performance are provided.

Keywords: Robust observers, learning-based estimation, parametric estimation, Gaussian process, linear matrix inequalities, data-driven methods.

1. Introduction

With modern dynamical systems growing in complexity, learning for state estimation has gained relevance and provides a suitable solution for problems with incomplete model knowledge, such as systems with unknown or unmodeled dynamics. Leveraging data generated from these systems and employing function approximators allows one to design estimators despite incomplete model descriptions by identifying the unknown dynamics. While the most common applications of learning via function approximators are in identifying dynamical systems from measurement data [1, 2], approximating control laws using offline data [3, 4], or adapting with data to generate optimal control policies online [5, 6], the utility of data and approximators for state estimation in nonlinear systems with unmodeled dynamics remains relatively unexplored.

Some early investigations into data-driven observers, for example, in [7] assume model availability. However, the current wave of data-driven control has demonstrated the effectiveness of approximators in controlling systems with partial models or, in some cases, a ‘model-free’ manner. Data-driven observers in the model-free setting were explored almost two decades ago in [8], where the authors proposed an adaptation rule for learning the weights of a linear-in-parameter neural network (LPNN) that results in uniformly ultimately bounded estimation error dynamics.

Although this work has been adopted in multiple applications such as robot control [9, 10], rotors [11], and more recently, wind turbines [12], the inherent assumptions and theory have hardly evolved. In most of these methodologies, the activation functions are considered to be radial basis functions, there is no measurement noise, and the theoretical guarantees of learning performance remain the same; an exception is [12] where the authors investigate input-to-state stability (ISS) observers for the known component of the model, but the learner performance is not ISS, and the learner’s weights require manual tuning.

In [13], the authors studied the problem of learning-based observer design for a class of partial differential equations (PDEs). The idea is based on designing a robust controller with respect to the structured model uncertainties, complemented with a learning layer, to auto-tune the observer gains, using a data-driven extremum seeking algorithm. However, the model uncertainties were compensated for by a robust estimation approach, and no online estimation of the model was proposed in this work. In [14], the authors proposed a learning-based observer design for the particular case of the Boussinesq PDE equations with parametric uncertainties, for a thermo-fluid application. The main idea is to construct a nominal observer which ensures ISS between the state estimation error and the parameter estimation error. The parameters are then estimated online using an extremum seeking method. The result was limited to the specific case of Boussinesq equations with structured parametric uncertainties.

In [15], the problem of adaptive observer design for a class of nonlinear systems was studied. The authors proposed to use shallow neural nets, with Lipschitz activation

*Corresponding author. Phone: +1 (617) 758-6175.

**A one-column draft format of this manuscript is available on ArXiv.

Email addresses: chakrabarty@merl.com (Ankush Chakrabarty), benosman@merl.com (Mouhacine Benosman)

functions, to estimate the unknown parts of the model. The use of Lipschitz constants of these activation functions simplifies the data-driven observer design procedure via new LMI conditions, ensuring pre-computable robust stability guarantees on the resulting state estimation error. The learning results, however, are model-based in the sense that the adaptation laws designed to learn the neural net coefficients are based on the model structure, similar to an indirect model-based adaptation approach, which limits the validity of the learning laws to a specific class of models; see for example [16] for a discussion about model-based and model-free learning and adaptation.

In [17], the authors propose to formulate the problem of observer design for a class of partially known linear and nonlinear systems as an optimal problem with quadratic cost over an infinite time support. Then the authors propose to use adaptive dynamic programming to solve this equivalent optimal control problem. However, this leads to a bounded integral cost of the estimation error, rather than, the usual sought for convergence result to a zero estimation error. This method also requires a specific structure on the output matrix that is not always available.

In [18], the authors propose, in the context of interval observers, a Gaussian process-based observer for a class of partially known nonlinear autonomous systems. The Gaussian process is used to approximate the unknown part of the model, which assumed to be a Lipschitz continuous function.

Contributions

In this paper, we extend the ideas from [14, 15] to the general case of partially known nonlinear models, where we propose to approximate the unknown part using neural networks, under a robustness design constraint in the form of local input-to-state stability between the state estimation error and the model parametric estimation error. This approach allows us to obtain a *safe* learning of the model, in the sense of maintaining boundedness of the state estimation error at all time, even during learning. After convergence of the learning algorithm, we propose to add a redesign step, which takes into account the learned model to improve the observer performance.

Organization

The rest of the paper is organized as follows. Section 2 describes the overall problem we are solving in this paper, along with assumptions. Section 3 introduces a method for generating initial robust observer gains, and a method for learning the unmodeled dynamics via Bayesian optimization is discussed in Section 4. An optional re-design of the observer gain is presented in Section 5 for improving the convergence rate by exploiting the learned nonlinearity, and the potential of our proposed approach is demonstrated using a numerical example in Section 6. Conclusions are drawn in Section 7.

Notation

We denote by \mathbb{R} the set of real numbers, \mathbb{R}_+ as the set of positive reals, and \mathbb{N} as the set of natural numbers. In a metric space (\mathbb{X}, ρ) , we denote an open ball centered at $x_0 \in \mathbb{X}$ with radius ε as $\mathbb{B}_\varepsilon(x_0) := \{x \in \mathbb{X} \mid \rho(x, x_0) < \varepsilon\}$. The symbols $\mathbf{E}(\cdot)$ and $\mathbf{V}(\cdot)$ denote the expectation and variance operators of a random variable, respectively. The notation \ln represents the natural logarithm. A function $f(x) \in \mathcal{C}^n$ if it is n -times differentiable and all its derivatives are continuous, and f is L -Lipschitz continuous (or L -Lipschitz, for brevity) if $\|f(x) - f(y)\| \leq L\|x - y\|$ for some $L > 0$ and any (x, y) pair in its domain. A function $f \in \mathcal{L}_\infty$ if $\|f(\cdot)\| < \infty$ on its domain. A continuous function $f : [0, \infty) \rightarrow [0, \infty)$ belongs to class \mathcal{K} if it is strictly increasing and $f(0) = 0$ and of class \mathcal{K}_∞ if it is of class \mathcal{K} and unbounded (with range extended to $[0, \infty]$). A continuous function f belongs to class \mathcal{KL} if, for each fixed t , the mapping $f(t, x)$ belongs to class \mathcal{K} with respect to t and, for each fixed x , the mapping $f(t, x)$ is decreasing with respect to t and $f(t, x) \rightarrow 0$ as $t \rightarrow \infty$. For a square matrix M , its largest and smallest eigenvalue are denoted by $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$, respectively. The transpose of a real-valued matrix M is denoted M^\top , and its Frobenius norm is denoted $\|M\|$. The quadratic form is shortened to be $\|x\|_Q^2 := x^\top Qx$.

2. Problem Statement

2.1. Preliminaries

We consider nonlinear systems modeled by

$$x_{t+1} = Ax_t + Bu_t + \phi(q_t), \quad (1a)$$

$$y_t = Cx_t, \quad (1b)$$

$$q_t = C_q x_t, \quad (1c)$$

where $t \in \mathbb{R}$ is the time index, $x \in \mathbb{R}^{n_x}$ denotes the system state, $u \in \mathbb{R}^{n_u}$ the known control input, $y \in \mathbb{R}^{n_y}$ the measured output, and $q_t \in \mathbb{Q} \subset \mathbb{R}^{n_q}$ the argument of the nonlinearity ϕ . The pair (A, B) is stabilizable, the pair (A, C) is observable, and the set of arguments \mathbb{Q} is compact. Our knowledge of the system is summarized in the following assumption.

Assumption 1. The system matrices A , B , and C are known. The nonlinearity's argument matrix C_q is known. The nonlinearity ϕ is completely unknown.

We also add a boundedness and Lipschitz assumption on the nonlinearity.

Assumption 2. The unknown nonlinearity is bounded: there exists a scalar $\bar{\phi} > 0$ such that $\phi(q) \leq \bar{\phi}$ for all $q \in \mathbb{Q}$. Furthermore, ϕ is Lipschitz continuous on \mathbb{Q} .

Assumption 3. The discrete Cauchy problem (1) admits a bounded¹ solution for any initial condition $x_0 \in \mathbb{R}^{n_x}$.

¹as defined in [19], Definition 1.1

Since every Lipschitz continuous nonlinearity is *a priori* continuous, and the domain of ϕ (namely, \mathbb{Q}) is compact by assumption, one can use a wide range of function approximators to express ϕ as a basis expansion. These approximators need only induce functions that are dense in the space of continuous functions such as polynomials (recall the Stone-Weierstrass Theorem) or ‘universal approximators’ like shallow/deep neural networks with appropriate activation functions [20, 21]. Therefore, we can rewrite the nonlinearity ϕ in the basis expansion form

$$\phi(q) = p_\star^\top \psi(q) + \epsilon_\phi(q), \quad (1d)$$

where p_\star denotes a $N_b \times n_x$ matrix of true coefficients of the basis expansion, ψ denotes a N_b -dimensional vector of \mathcal{C}^1 basis/activation functions, and ϵ_ϕ denotes a bounded approximation error; that is, $\|\epsilon_\phi(q)\| \leq \bar{\epsilon}_\phi$ for any $q \in \mathbb{Q}$. This basis function form is also referred to in the literature as a neural approximator [17]. We note that the \mathcal{C}^1 nature of ψ implies that it is Lipschitz continuous on \mathbb{Q} ; we denote its Lipschitz constant as \mathfrak{L}_ψ . Note that (1d), together with Assumption 2 and the boundedness of p_\star on \mathbb{Q} , implies boundedness of ψ on \mathbb{Q} .

Assumption 4. There exists a known positive scalar \bar{p}_\star such that $\|p_\star\| \leq \bar{p}_\star$.

Assumption 4 is admittedly strong, but if such a \bar{p}_\star is not known exactly, a good estimate of \bar{p}_\star is rarely needed in practice, as we shall see in the remainder of the paper. In fact, this assumption mainly ensures good robustness properties of the initial observer. As data is obtained, the learning algorithm will overcome the conservativeness of the estimate and \bar{p}_\star will be computed (not guessed) for the redesign.

The system (1) motivates a state observer of the form

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + p_t^\top \psi(\hat{q}_t) + L_0(C\hat{x}_t - y_t), \quad (2a)$$

$$\hat{q}_t = C_q \hat{x}_t, \quad (2b)$$

with gain matrix $L_0 \in \mathbb{R}^{n_x \times n_y}$. The error dynamics of the observer (2) with $e_t := \hat{x}_t - x_t$ and $e_t^p = p_t - p_\star$ is given by

$$e_{t+1} = A_L e_t + \psi(\hat{q}_t)^\top e_t^p + p_\star^\top \Delta \psi_t - \epsilon_\phi(q_t), \quad (3)$$

where $A_L = A + L_0 C$ and $\Delta \psi_t = \psi(\hat{q}_t) - \psi(q_t)$. Our initial step is to design a learning algorithm and the observer gain L_0 such that

$$e_t \rightarrow \mathbb{B}_\varepsilon(0) \text{ and } p_t \rightarrow p_\star \text{ as } t \rightarrow \infty,$$

and then switch to a more precise observer, once we have learned the model uncertainties.

To reiterate, our data-driven observer operates in three design modes: (i) in the data collection mode, we design L_0 such that the error dynamics are locally input-state stable (L-ISS) between the parameter estimation error $p_t^\top - p_\star$ and the state estimation error e_t , setting the

stage for safe learning; (ii) in the safe learning mode, we leverage Bayesian optimization to learn the coefficients p_t ; and, (iii) as the learning terminates, we redesign L_0 taking the learned nonlinearity into account to improve the performance of the observer, based on the learned model.

3. Nominal ISS observer design

The main idea of the nominal observer is to design the gain L_0 such that one can guarantee a local ISS property between the parameter estimation error and the state estimation error. In this section, we present conditions that, if satisfied, enable the design of a suitable L_0 . Consequently, we propose a simple convex programming formulation for computing an L_0 that satisfies these conditions.

3.1. Conditions for initial design

We begin with the following definitions from [22].

Definition 1 (L-ISS). A discrete-time nonlinear system

$$\xi_{t+1} = \varphi(\xi_t, \nu_t) \quad (4)$$

is locally input-to-state stable (L-ISS) with respect to the exogenous input ν_t , if there exist scalars $\varrho_\xi > 0$, $\varrho_\nu > 0$, $\gamma \in \mathcal{K}_\infty$ and a function $\beta(\cdot, \cdot) \in \mathcal{KL}$ such that for all $\|x_0\| \leq \varrho_\xi$ and $\|\nu\|_\infty \leq \varrho_\nu$, the condition

$$\|\xi(t, \xi_0, \nu)\| \leq \beta(t, \|\xi_0\|) + \gamma(\|\nu\|_\infty)$$

is satisfied, for all $t \in \mathbb{N}$.

Definition 2 (L-ISS Lyapunov function). Let the set $\mathcal{G} \subset \mathbb{R}^n$ contain the origin in its interior. A Lipschitz continuous function $V : \mathcal{G} \rightarrow [0, \infty)$ is called a local ISS Lyapunov function for a system (4) on \mathcal{G} if there exist \mathcal{K}_∞ functions $\gamma_1(\cdot)$, $\gamma_2(\cdot)$, $\gamma(\cdot)$ and $\beta(\cdot)$ such that

$$\gamma_1(\|\xi\|) \leq V(\xi) \leq \gamma_2(\|\xi\|) \quad (5a)$$

$$V(\varphi(\xi, \nu)) - V(\xi) \leq -\gamma(\|\xi\|) + \beta(\|\nu\|) \quad (5b)$$

is satisfied, for all $\xi \in \mathcal{G}$ and $\nu \in \Xi$.

Lemma 1. *If the system (4) admits a locally ISS Lyapunov function, then it is locally ISS w.r.t. ν .*

Proof. See [23, Theorem 4]. \square

The following theorem provides conditions for designing L_0 such that $(A + L_0 C)$ is Schur (stable), and the state estimation error e is L-ISS with respect to the parameter estimation error e^p and the neural approximation error ϵ_ϕ .

Theorem 1. *Recall $A_L = A + L_0 C$. Let $\mathfrak{L}_\psi := \max_{\mathbb{Q}} \|\nabla_q \psi\|$ and Assumptions 1–4 hold. Suppose there exist matrices $P = P^\top > 0$, $Q = Q^\top > 0$, and an initial observer gain L_0 such that*

$$A_L^\top P A_L - P + Q = 0, \quad (6a)$$

$$4\lambda_{\max}(P)\bar{p}_\star^2 \mathfrak{L}_\psi^2 \|C_q\|^2 + 8\bar{p}_\star \mathfrak{L}_\psi \|P A_L\| \|C_q\| \leq \lambda_{\min}(Q) \quad (6b)$$

then the error system (3) is L-ISS w.r.t. $\begin{bmatrix} e_t^p, \epsilon_\phi \end{bmatrix}^\top$.

Proof. We consider a candidate of L-ISS Lyapunov function V defined by

$$V_t = e_t^\top P e_t, \quad (7)$$

For quadratic V , the condition (5a) is satisfied with $\gamma_1 = \lambda_{\min}(P)\|e\|^2$ and $\gamma_2 = \lambda_{\max}(P)\|e\|^2$. Taking the time difference $\Delta V_t := V_{t+1} - V_t$ of (7) along the solution of (3), we obtain

$$\Delta V_t = \sum_{\ell=1}^6 v_\ell, \quad (8)$$

where

$$\begin{aligned} v_1 &= e_t^\top (A_L^\top P A_L - P) e_t = -e_t^\top Q e_t, \\ v_2 &= (p_\star^\top \Delta \psi_t + \psi(\hat{q})^\top e_t^p)^\top P (p_\star^\top \Delta \psi_t + \psi(\hat{q})^\top e_t^p), \\ v_3 &= \epsilon_\phi^\top P \epsilon_\phi, \\ v_4 &= 2(A_L e_t)^\top P (p_\star^\top \Delta \psi_t + \psi(\hat{q})^\top e_t^p), \\ v_5 &= -2\epsilon_\phi^\top P A_L e_t, \\ v_6 &= -2\epsilon_\phi^\top P (p_\star^\top \Delta \psi_t + \psi(\hat{q})^\top e_t^p). \end{aligned}$$

where the rightmost equation of v_1 is obtained by replacing with Q using (6a).

Before further analysis, note that we can bound the norms of $p_\star^\top \Delta \psi_t$ and $\psi(\hat{q})^\top e_t^p$ as follows:

$$\begin{aligned} \|p_\star^\top \Delta \psi_t\| &\leq \bar{p}_\star \mathfrak{L}_\psi \|q_t - \hat{q}_t\| \leq \bar{p}_\star \mathfrak{L}_\psi \|C_q\| \|e_t\|, \\ \|\psi(\hat{q})^\top e_t^p\| &\leq \|\psi(\hat{q})\| \|e_t^p\| = \bar{\psi} \|e_t^p\|, \end{aligned}$$

where the first bound is a consequence of the Lipschitz property of ψ and the second due to \mathbb{Q} being compact and ψ being bounded on \mathbb{Q} . Using these bounds, we can proceed with bounded the v_ℓ terms. Concretely,

$$\begin{aligned} v_1 &\leq -\lambda_{\min}(Q) \|e_t\|^2, \\ v_2 &\leq \lambda_{\max}(P) (\bar{p}_\star^2 \mathfrak{L}_\psi^2 \|C_q\|^2 \|e_t\|^2 + \bar{\psi}^2 \|e_t^p\|^2), \\ v_3 &\leq \lambda_{\max}(P) \|\epsilon_\phi\|^2, \\ v_4 &\leq 2\bar{p}_\star \mathfrak{L}_\psi \|P A_L\| \|C_q\| \|e_t\|^2 + 2\bar{\psi} \|P A_L\| \|e_t^p\| \|e_t\|, \\ v_5 &\leq 2\bar{\epsilon}_\phi \|P A_L\| \|e_t\|, \\ v_6 &\leq 2\bar{p}_\star \bar{\epsilon}_\phi \mathfrak{L}_\psi \|P\| \|C_q\| \|e_t\| + 2\bar{\psi} \|e_t^p\|. \end{aligned}$$

Collecting relevant terms, we can write

$$\Delta V_t \leq -a_0 \|e_t\|^2 + a_1 \|e_t\|^2 + a_2 \|e_t\| + a_3 \|e_t^p\|^2 + a_4 \|\epsilon_\phi\|^2,$$

where

$$\begin{aligned} a_0 &= \frac{1}{2} \lambda_{\min}(Q) > 0, \\ a_1 &= -\frac{1}{4} \lambda_{\min}(Q) + \lambda_{\max}(P) \bar{p}_\star^2 \mathfrak{L}_\psi^2 \|C_q\|^2 + 2\bar{p}_\star \mathfrak{L}_\psi \|P A_L\| \|C_q\|, \\ a_2 &= -\frac{1}{4} \lambda_{\min}(Q) \|e_t\| + (2\|P A_L\| + 2\bar{p}_\star \mathfrak{L}_\psi \|P\| \|C_q\|) \bar{\epsilon}_\phi \\ &\quad + 2\bar{\psi} \|P A_L\| \|e_t^p\|, \\ a_3 &= \lambda_{\max}(P) \bar{\psi}^2 > 0, \\ a_4 &= \lambda_{\max}(P) > 0. \end{aligned}$$

Since the observer gain L_0 satisfies the condition (6b), we get $a_1 \leq 0$. Therefore, if $a_2 \leq 0$, that is, if

$$\|e_t^p\| \leq \frac{\lambda_{\min}(Q)}{8\bar{\psi} \|P A_L\|} \|e_t\| - \frac{\|P A_L\| + \bar{p}_\star \mathfrak{L}_\psi \|P\| \|C_q\|}{\bar{\psi} \|P A_L\|} \|\epsilon_\phi\|,$$

then

$$\Delta V_t \leq -a_0 \|e_t\|^2 + a_3 \|e_t^p\|^2 + a_4 \|\epsilon_\phi\|^2, \quad (9)$$

which satisfies the condition (5b) with ν replaced by $\begin{bmatrix} e_t^p, \epsilon_\phi \end{bmatrix}^\top$. \square

3.2. Computing the initial observer gain

We begin with the following theorem that provides a design procedure for L_0 for a fixed \mathfrak{L}_ψ . Such an \mathfrak{L}_ψ could be known, or, if unknown, one could design an observer for a maximal \mathfrak{L}_ψ (obtained via a line search) such that the conditions (6) are satisfied. The following convex relaxation is proposed.

Theorem 2. For fixed \bar{p}_\star and \mathfrak{L}_ϕ , if there exist matrices $P = P^\top$, $Q = Q^\top$, $K \in \mathbb{R}^{n_x \times n_y}$, and scalars $\beta_\kappa > 0$, $\kappa_1 > 0$, $\kappa_2 > 0$, and $\kappa_3 > 0$ such that

$$\min_{P, K, Q, \kappa_0, \kappa_1, \kappa_2, \kappa_3} \kappa_0 + \beta_\kappa (\kappa_1 + \kappa_2 + \kappa_3) \quad (10a)$$

subject to:

$$\begin{bmatrix} -P + Q + \kappa_0 I & \star \\ P A + K C & P \end{bmatrix} \leq 0 \quad (10b)$$

$$Q \geq \kappa_1 I \quad (10c)$$

$$0 < P \leq \kappa_2 I \quad (10d)$$

$$\|P A + K C\| \leq \kappa_3 \quad (10e)$$

$$4\bar{p}_\star^2 \mathfrak{L}_\psi^2 \|C_q\|^2 \kappa_2 + 8\bar{p}_\star \mathfrak{L}_\psi \|C_q\| \kappa_3 - \kappa_1 \leq 0 \quad (10f)$$

has an optimal solution with $\kappa_0 = 0$, then an observer of the form (2) with gain $L_0 = P^{-1}K$ yields error dynamics (3) that are L-ISS.

Proof. We begin by noting $K := P L_0$ and substituting K into (6). This yields

$$(P A + K C)^\top P^{-1} (P A + K C) - P + Q = 0, \quad (11)$$

$$4\lambda_{\max}(P) \bar{p}_\star^2 \mathfrak{L}_\psi^2 \|C_q\|^2 + 8\bar{p}_\star \mathfrak{L}_\psi \|P A + K C\| \|C_q\| \leq \lambda_{\min}(Q). \quad (12)$$

We relax the equality (11) with the inequality

$$(P A + K C)^\top P^{-1} (P A + K C) - P + Q \leq -\kappa_0 I. \quad (13)$$

Taking Schur complements of (13) yields (10b). Conditions (10c)–(10e) provide bounds on $\lambda_{\min}(Q)$, $\lambda_{\max}(P)$, and $\|PA + KC\|$: with these bounds κ_1 , κ_2 , and κ_3 , one can explicitly write the constraint (12) as (10f). If the optimal value of $\kappa_0 = 0$, then the equality (11) is exact, and the conditions in (6) are satisfied. \square

Remark 1. Note that, for a fixed \mathfrak{L}_ψ and \bar{p}_\star , the problem (10) is convex, and therefore can be solved efficiently using standard convex solvers such as CVX/YALMIP.

Remark 2. The scalar β_κ is a regularization parameter that should be kept small enough to promote the computation of unique solutions while ensuring the focus of the objective function is to force $\kappa \approx 0$.

Remark 3. One can use the conditions in (10) to perform a grid search for the largest Lipschitz constant and the largest coefficient bound for which a feasible solution to (6) exists. Concretely, by solving the problem

$$(\hat{\mathfrak{L}}_\psi, \hat{p}_\star) := \arg \max(\mathfrak{L}_\psi + \bar{p}_\star) \text{ subject to: (10)} \quad (14)$$

and ensuring $\kappa_0 = 0$ for the optimal solution of (14) generates $\hat{\mathfrak{L}}_\psi$ and \hat{p}_\star and L_0 such that the error dynamics (3) are L-ISS, as long as these are overestimates; i.e. $\hat{\mathfrak{L}}_\psi \geq \mathfrak{L}_\psi$ and $\hat{p}_\star \geq \bar{p}_\star$. Of course, if either one is known *a priori*, then one can perform a line search for the other.

4. Learning via Bayesian optimization

The previous section described a method for starting with a *safe* initial state estimator: one whose error trajectory does not diverge despite learning, as long as the coefficients p_t of the parametric model are not unbounded. In this section, we provide a data-driven method for updating the model from on-line data using Bayesian optimization.

4.1. Data collection and reward

Let T_ℓ denote the horizon over which measurements and state estimates are collected, and p_0 denote the initial guess for the model coefficients. We begin with no data, and an initial observer gain L_0 . We run the observer (2) over the learning horizon $\{0, 1, \dots, T_\ell - 1\}$ and compute the reward

$$\mathcal{J}(p_j) := - \left(\sum_{t=0}^{T_\ell-1} \|C\hat{x}_t - y_t\|_{W_1}^2 + \frac{1}{T_\ell} \|p_j\|_{W_2}^2 \right) \quad (15)$$

with reward weighting matrices $W_1 > 0$ and $W_2 \geq 0$. The objective of the second term in this reward function is to promote a unique solution by regularizing the learned coefficients; typically W_2 is small. From this batch of measurements, we get a data sample $(p_0, \mathcal{J}(p_0))$. With the j -th iteration dataset

$$\mathcal{D}_j = (p_{0:j-1}, \mathcal{J}(p_{0:j-1})),$$

we use Bayesian optimization to compute a solution to the problem

$$p_\infty = \arg \max_{p \in \mathbb{P}} \mathcal{J}(p), \quad (16)$$

where \mathbb{P} is an admissible compact and convex set of parameters.

Since the gradient of \mathcal{J} cannot be evaluated analytically, and the data could have inherent noise, one cannot use standard gradient-based tools to compute p_∞ . Instead, we resort to learning a surrogate model of \mathcal{J} using Gaussian process (GP) regression [24], and exploiting the statistics of the learned surrogate to inform exploration and exploitation; that is, how to choose p_j based on the current dataset \mathcal{D}_{j-1} . In particular, we use the expected-improvement (EI) method for acquiring subsequent p_j values. The combination of GP modeling and EI acquisition is referred to herein as GP-EI, and is the topic of the next subsection.

4.2. Bayesian Optimization with the GP-EI algorithm

We formally reiterate our assumptions on \mathcal{J} and p_\star in the following assumption.

Assumption 5. The reward function \mathcal{J} is continuous with respect to its argument p for every $p \in \mathbb{P}$. Furthermore, the function has a unique global maximizer p_\star on the set \mathbb{P} .

Since \mathcal{J} is assumed continuous, we leverage the data at the j -th iteration to construct a surrogate GP model of the reward, given by

$$\hat{\mathcal{J}}_j := \text{GP}(\mu(p; \mathcal{D}_j), \sigma(p, p'; \mathcal{D}_j)), \quad (17)$$

where $\mu(\cdot)$ is the predictive mean function, and $\sigma(\cdot, \cdot)$ is the predictive variance function. Typically, the variance is expressed through the use of kernels. A commonly used kernel is the squared exponential (SE) kernel

$$\mathcal{K}_{\text{SE}} = \sigma_0^2 \exp\left(-\frac{1}{2}r^2\right), \quad (18)$$

with

$$r^2 \equiv r^2(p, p') = \frac{\|p - p'\|^2}{\sigma_1^2}$$

and hyperparameters σ_0 (the output variance) and σ_1 (the length scale). However, the SE kernel sometimes results in overly smooth functions; to avoid this, another class of kernels called Matérn kernels have gained popularity, of which the Matérn 5/2 kernel has the form

$$\mathcal{K}_{\text{M5/2}} = \sigma_0 \left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) \exp\left(-\sqrt{5}r\right). \quad (19)$$

At the j -th learning iteration, for a new query sample $p \in \mathbb{P}$, the GP model predicts the mean and variance of the

reward to be

$$\begin{aligned}\mu(p) &= k_j(p)^\top K_{j-1}^{-1} \mathcal{J}_{0:j-1} \\ \sigma(p) &= \mathcal{K}(p, p) - k_j(p) K_{j-1}^{-1} k_j(p)^\top,\end{aligned}$$

where

$$\begin{aligned}k_j(p) &= [\mathcal{K}(p_0, p) \quad \mathcal{K}(p_1, p) \quad \cdots \quad \mathcal{K}(p_{j-1}, p)], \\ K_{j-1} &= \begin{bmatrix} \mathcal{K}(p_0, p_0) & \cdots & \mathcal{K}(p_0, p_{j-1}) \\ \vdots & \ddots & \vdots \\ \mathcal{K}(p_{j-1}, p_0) & \cdots & \mathcal{K}(p_{j-1}, p_{j-1}) \end{bmatrix}.\end{aligned}$$

In Bayesian optimization, we use the mean and variance of the surrogate model $\hat{\mathcal{J}}_j$ in (17) to construct an acquisition function to inform the selection of a p_j that maximizes the likelihood of improving the current best reward

$$\hat{\mathcal{J}}_j^* := \max_{p \in \mathbb{P}} \hat{\mathcal{J}}_j(p).$$

To this end, we define an improvement function

$$\mathcal{I} := \mathcal{I}(p, j) = \max\{0, \mathcal{J}(p) - \hat{\mathcal{J}}_j^*\},$$

whose likelihood function, based on a Gaussian posterior distribution $\mathcal{N}(\mu, \sigma^2)$, is given by

$$\mathbb{L}(\mathcal{I}) = \frac{1}{\sqrt{2\pi}\sigma(p)} \exp\left(-\frac{1}{2} \frac{(\mu(p) - \hat{\mathcal{J}}_j^* - \mathcal{I})^2}{\sigma^2(p)}\right).$$

Taking an expectation of the likelihood function yields

$$\begin{aligned}\mathbb{E}(\mathcal{I}(p, j)) &= \mathbb{E}(\mathcal{I}(p, j)) \\ &= \int_0^\infty \frac{\mathcal{I}}{\sqrt{2\pi}\sigma(p)} \exp\left(-\frac{1}{2} \frac{(\mu(p) - \hat{\mathcal{J}}_j^* - \mathcal{I})^2}{\sigma^2(p)}\right) d\mathcal{I}.\end{aligned}$$

Performing a change of variables and integrating by parts yields

$$\mathbb{E}(\mathcal{I}(p, j)) = \begin{cases} \sigma(p)\gamma(z) + (\mu(p) - \hat{\mathcal{J}}_j^*)\Gamma(z), & \text{if } \sigma(p) > 0, \\ 0 & \text{if } \sigma(p) = 0. \end{cases}$$

where

$$z = \frac{\mu(p) - \hat{\mathcal{J}}_j^*}{\sigma(p)},$$

and $\gamma(\cdot)$, $\Gamma(\cdot)$ are the PDF and the CDF of the mean-zero unit-variance normal distribution, respectively.

In the j -th iteration of learning, we use the data \mathcal{D}_j to construct the EI acquisition function using the surrogate $\hat{\mathcal{J}}_j$. Subsequently, we sample on \mathbb{P} and obtain

$$p_j = \arg \max_{p \in \mathbb{P}} \mathbb{E}(\mathcal{I}(p, j)), \quad (20)$$

which serves as the estimate of the model coefficients p_t

in (2) until the next learning iteration. We terminate the learning algorithm when

$$\mathbb{E}(\mathcal{I}(p, j)) < \varepsilon_{\text{EI}}, \quad \forall p \in \mathbb{P} \quad (21)$$

for some small threshold $\varepsilon_{\text{EI}} > 0$. The terminal set of coefficients is defined p_∞ .

Remark 4. We do not explicitly write a hyperparameter selection procedure as it is beyond the scope of this work. Standard methods such as log-marginal-likelihood maximization is used for implementation purposes to find good GP variances and length scales. For more details, we refer the interested reader to [24, Chapter 5].

4.3. Regret analysis

In this subsection, we quantify the performance of the GP-EI learning algorithm in terms of the cumulative regret

$$\mathcal{R}_N := \sum_{j=0}^N \mathcal{J}(p_\star) - \mathcal{J}(p_j), \quad (22)$$

where $N \in \mathbb{N}$ denotes the number of learning/training iterations. Specifically, we will demonstrate that the regret associated with GP-EI in our learning-based observer is sublinear, and therefore, $\mathcal{R}_N/N \rightarrow 0$ as $N \rightarrow \infty$.

Proposition 1 ([25]). *Suppose Assumption 5 holds. Let $\delta \in (0, 1)$ and ε_{EI} be a fixed termination threshold. Assume that \mathcal{J} lies in the reproducible kernel Hilbert space $\mathbb{H}_{\mathcal{K}}(\mathbb{P})$ corresponding to the kernel $\mathcal{K}(p, p')$, and that the noise corrupting the reward function has zero mean conditioned on the noise history, and is bounded almost surely. Let $\|\mathcal{J}\|_{\mathcal{K}}^2 \leq B_{\mathcal{J}}$ and $\zeta_t := 2B_{\mathcal{J}} + 300\chi_t \log^3(t/\delta)$, where χ_t is a kernel-dependent constant depending on t . Then the GP-EI algorithm with termination criterion (21) has a probabilistic bound*

$$\Pr\left(\mathcal{R}_N \leq \sqrt{N\zeta_N\chi_N}\right) \geq 1 - \delta \quad (23)$$

on the cumulative regret over N learning iterations.

Using the results of Theorem 1 and Proposition 1, under Assumptions 1–5, it is easy to write that the estimation error bound satisfies

$$\|e_t(\xi_0, \nu)\| \leq \beta(t, \|\xi_0\|) + \gamma(\|\nu(t)\|_\infty), \quad (24)$$

where $\nu(t) = [e_t^{p^\top}, \epsilon_\phi^\top]^\top$, and e_t^p satisfies the probabilistic bound of Proposition 1.

Except for the case where the learning cost is assumed to be strongly locally convex in a neighborhood of p^\star , the regret bound (23) in general, does not allow us to write an explicit bound on the parameter estimation error e^p . However, due to the continuity of the \mathcal{R}_N in the variable e^p , one can use classical arguments from transformation theory of random variables, e.g. ([26], p. 68), to conclude about the boundedness of the parameter estimation error

e^p . This, together with (24), allows us to guarantee the safety, in the form of boundedness of the state estimation error e , of the observer during the learning iterations.

Remark 5. For squared exponential kernels, the constant χ_N is of the order $(\log(T))^{n_p+1}$, where n_p denotes the dimensionality of \mathbb{P} . For 5/2 Matérn kernels, χ_N is of the order $T^{\frac{n_p(n_p+1)}{5+n_p(n_p+1)}} \log(T)$. For both, the cumulative regret is sublinear as per the bound in Proposition 1.

5. Observer gain redesign

Upon completion of the Bayesian optimization stage, we have learned the unmodeled components of the dynamical system to some accuracy. Therefore, one can (optionally) leverage this newly acquired model knowledge to update the observer gains from the initial design.

Let p_∞ denote the value of p_t which minimizes simple regret over all learning iterations. We propose a redesigned observer of the form

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + p_\infty\psi(\hat{q}_t) + L(C\hat{x}_t - y_t), \quad (25)$$

where the redesigned observer gain is L . The error dynamics of this redesigned observer can be rewritten as

$$e_{t+1} = (A + LC)e_t + p_\infty^\top \Delta\psi + \psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t), \quad (26)$$

where $e_\infty^p = p_\infty - p_*$ is a constant parameter estimation error. In this case, the estimated value p_∞ will be used to redesign the observer gains. The following theorem encapsulates the redesign conditions.

Theorem 3. *If there exist matrices K , $P = P^\top > 0$, and $Q = Q^\top > 0$ such that*

$$\begin{bmatrix} -P + Q + \mathfrak{L}_\psi^2 C_q^\top C_q & * & * \\ PA + KC & -P & 0 \\ (PA + KC)p_\infty^\top & 0 & p_\infty P p_\infty^\top - I \end{bmatrix} \leq 0, \quad (27)$$

then the redesigned observer (25) with gain $L = P^{-1}K$ makes the error dynamics (3) L -ISS with respect to $[e_\infty^p, \epsilon_\phi]^\top$.

Proof. Let $V_t = e_t^\top P e_t$ and replace $PA + KC$ by $P(A + LC)$ by substituting $L = P^{-1}K$. Then taking successive Schur complements of the inequality in (27) yields

$$\begin{bmatrix} \Omega + Q & * \\ P(A + LC)p_\infty^\top & p_\infty P p_\infty^\top - I \end{bmatrix} + \begin{bmatrix} \mathfrak{L}_\psi^2 C_q^\top C_q & 0 \\ 0 & -I \end{bmatrix} \leq 0, \quad (28)$$

with

$$\Omega = (A + LC)^\top P (A + LC) - P.$$

Taking a congruence transform with $[e_t^\top \quad \Delta\psi^\top]^\top$ yields

$$e_t^\top \Omega e_t + 2e_t^\top (A + LC)^\top P p_\infty^\top \Delta\psi + \Delta\psi^\top p_\infty P p_\infty^\top \Delta\psi + \mathfrak{L}_\psi^2 \Delta q^\top \Delta q - \Delta\psi^\top \Delta\psi \leq -e_t^\top Q e_t.$$

Since \mathfrak{L}_ψ is a Lipschitz constant of ψ , the above inequality implies that

$$\tilde{V} \leq -e_t^\top Q e_t, \quad (29)$$

where

$$\tilde{V} := e_t^\top \Omega e_t + 2e_t^\top (A + LC)^\top P p_\infty^\top \Delta\psi + \Delta\psi^\top p_\infty P p_\infty^\top \Delta\psi.$$

Let $V_t = e_t^\top P e_t$. Then,

$$\Delta V_t = \tilde{V} + \sum_{\ell=1}^3 \tilde{v}_\ell,$$

where

$$\begin{aligned} \tilde{v}_1 &= (\psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t))^\top P (\psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t)), \\ \tilde{v}_2 &= 2e_t^\top (A + LC)^\top P (\psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t)), \\ \tilde{v}_3 &= 2\Delta\psi^\top p_\infty P (\psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t)). \end{aligned}$$

Following similar arguments as the proof of Theorem 1, we can bound these terms as

$$\begin{aligned} \tilde{v}_1 &\leq \lambda_{\max}(P) \|\psi(q_t)^\top e_\infty^p - \epsilon_\phi(q_t)\|^2 \\ &\leq \lambda_{\max}(P) (\bar{\psi}^2 \|e_\infty^p\|^2 + \|\epsilon_\phi\|^2), \\ \tilde{v}_2 &\leq 2\|P(A + LC)\| \|e_t\| (\bar{\psi} \|e_\infty^p\| + \|\epsilon_\phi\|), \\ \tilde{v}_3 &\leq 2p_\infty \|P\| \|C_q\| \mathfrak{L}_\psi \|e_t\| (\bar{\psi} \|e_\infty^p\| + \|\epsilon_\phi\|). \end{aligned}$$

These bounds, along with (29), yields

$$\Delta V_t \leq -\delta_0 \|e_t\|^2 + \delta_1 \|e_\infty^p\|^2 + \delta_2 \|\epsilon_\phi\|^2 + \delta_3 \|e_t\|$$

where $\delta_0 = \frac{1}{2}\lambda_{\min}(Q)$, $\delta_1 = \lambda_{\max}(P)\bar{\psi}^2$, and $\delta_2 = \lambda_{\max}(P)$, and $\delta_3 \leq 0$ if

$$\|e_\infty^p\| \leq \frac{\lambda_{\min}(Q)}{4\bar{\psi} (\|P(A + LC)\| + p_\infty \mathfrak{L}_\psi \|P\| \|C_q\|)} \|e_t\| - \frac{1}{\bar{\psi}} \|\epsilon_\phi\|.$$

The rest of the proof is identical to the proof of Theorem 1. \square

Remark 6. Taking advantage of the learned coefficients p_∞ , we obtain simpler LMI conditions in (27). Furthermore, from the L-ISS definition, we deduce that redesigning the observer with a smaller parameter estimation error e_∞^p leads to a smaller state estimation error.

A complete pseudocode is provided in the Appendix.

6. Numerical Example

Consider the van der Pol oscillator system

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix} x_1^2 x_2, \quad y = x_1.$$

We begin by Euler discretization of the continuous-time dynamics with a sampling time of $\tau = 0.01$ s. Comparing

with (1), we have

$$A = \begin{bmatrix} 1 & \tau \\ \tau & 1 - \tau \end{bmatrix}, B = \begin{bmatrix} 0 \\ -\tau \end{bmatrix}, C = [1 \ 0], C_q = I,$$

and $\phi(q) = q_1^2 q_2$. We know A , B , C_q , and C , as per Assumption 1. Owing to limit cycle behaviour of the oscillator, Assumption 3 holds. Furthermore, the nonlinearity ϕ is locally Lipschitz, with unknown Lipschitz constant. Data is collected by using forward simulations of the oscillator system $T_\ell = 40$ s from an initial condition $[1, 1]^\top$.

We design the initial observer gain by solving the problem (14), assuming $\bar{p}_\star = 10^{-2}$. We run a golden-section search for $\mathfrak{L}_\phi \in [0, 10]$, which yields an initial observer gain of

$$L_0 = [1.1727 \ 7.3679]^\top$$

for $\hat{\mathfrak{L}}_\psi = 4.332$. The observer initial conditions are set to zero.

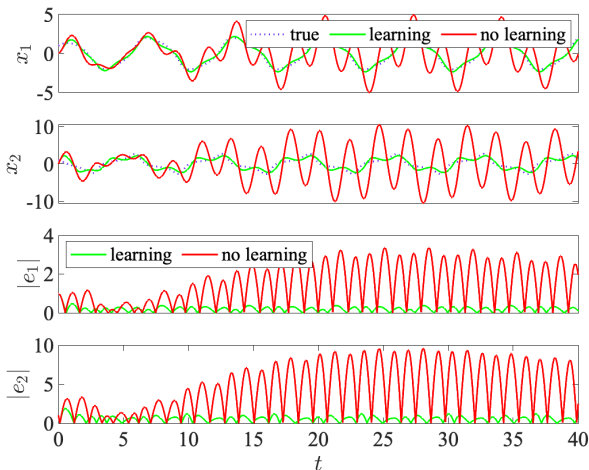


Figure 1: Comparison of state estimates and error norms with and without learning from data.

For fairness, we choose Legendre polynomials as basis functions and not monomials, which would perfectly fit the nonlinearity $\phi(q)$ with very few coefficients. Our basis functions are

$$\psi(\hat{q}) = 10 \begin{bmatrix} (3\hat{q}_1^2 - 1)(3\hat{q}_2^2 - 1) \\ (3\hat{q}_1^2 - 1)\hat{q}_2 \\ \hat{q}_1(3\hat{q}_2^2 - 1) \\ (5\hat{q}_1^3 - 3\hat{q}_1) \\ (5\hat{q}_2^3 - 3\hat{q}_2) \end{bmatrix},$$

where the scaling helps find good coefficients despite the strong bound on \bar{p}_\star . Performing Bayesian optimization with the GP-EI method using a weighting matrix $W_1 =$

200, $W_2 = 1$, and termination threshold $\varepsilon_{\text{EI}} = 0.01$ yields

$$p_\infty = 10^{-3} \times \begin{bmatrix} -0.6077 \\ 8.4930 \\ -9.2877 \\ 1.8897 \\ 9.8417 \end{bmatrix}$$

in $N = 200$ training iterations, with 1000 uniformly random samples on $[-\bar{p}_\star, \bar{p}_\star]^5$ to select p_j at each j using (20). The surrogate GP model is constructed at each time using sklearn in Python or MATLAB's `fitrgp` function in the Statistics and Machine Learning Toolbox. We use the Matérn ARD-5/2 kernel and select hyperparameters using BFGS on the log-likelihood function at each training iteration. The effects of model learning is illustrated in Figure 1. Prior to learning, the initial observer design provides bounded error dynamics (red continuous line), but the error is considerably reduced after learning is completed (green continuous line).

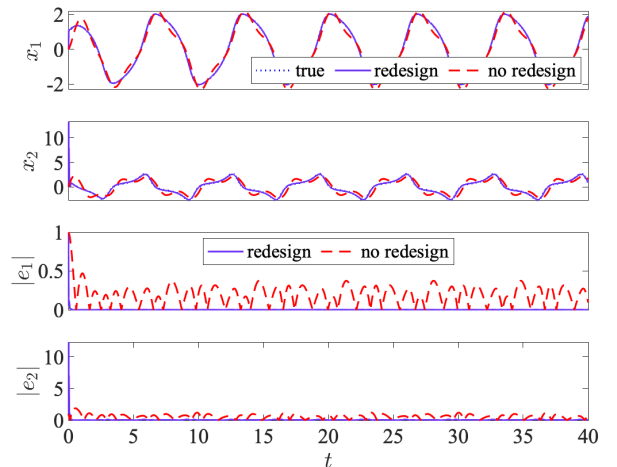


Figure 2: Comparison of state estimates and error norms with and without final redesign of the observer gains.

With the learned nonlinearity in a neural approximator/basis expansion form $p_\infty^\top \psi(\hat{q})$, we select $\mathbb{Q} = [-5, 5]^2$ and use the kernelized Lipschitz estimation method in [27] to compute $\hat{\mathfrak{L}}_\psi = 24.537$, with which we can solve (27) to get $L = [39.3221 \ 958.7488]^\top$. As seen from Figure 2, the redesign results in further reduction of state estimation error.

In order to demonstrate the effectiveness of the proposed algorithm, we compare the performance of our learning-based observer with well-known robust and adaptive observers in the literature. The observers we compare against include: the robust adaptive observer (RA) described in [28, Chapter 5], the high-gain (HG) observer described in [29], and the recurrent neural network (RNN) learning-based observer proposed in [30]. In all cases, we make multiple simplifications regarding knowledge of the model dynamics, since these observers failed to converge with

completely unknown dynamics. In the first two observers, one requires a ‘nominal model’ for $\phi(q)$, which we assume to be $\hat{x}_1^2 \hat{x}_2 + 0.2 * \sin(\hat{x}_2)$ to emulate a bounded disturbance on the model knowledge. For the neural observer, we start with an initial condition $[0.5 \ 0.5]^\top$ which is close to the true system state, because starting from the origin results in the neural approximator diverging. A comparison of the state estimation error in the unmeasured state, that is, $|e_2|$ is illustrated in Fig. 3. Clearly, our proposed observer (purple line) outperforms the others despite the simplifications made to help the other observer architectures. The RA observer (amber line) demonstrates good performance, but slow convergence rate, despite extensive hand-tuning - another drawback of this observer is that the observer state-space 8-dimensional and there are multiple hyperparameters and gains that need hand-tuning. The HG observer (green line), expectedly, has high state estimation error initially, although the steady-state estimation norm is smaller than RA and RNN, albeit larger than our proposed observer. Finally, the RNN observer (magenta line) exhibits significantly worse performance with time, and sometimes the estimation error grows large; this is because the observer estimation error dynamics are uniformly ultimately bounded, but not ISS, so large approximation errors in the recurrent neural network result in poor quality of the state estimates.

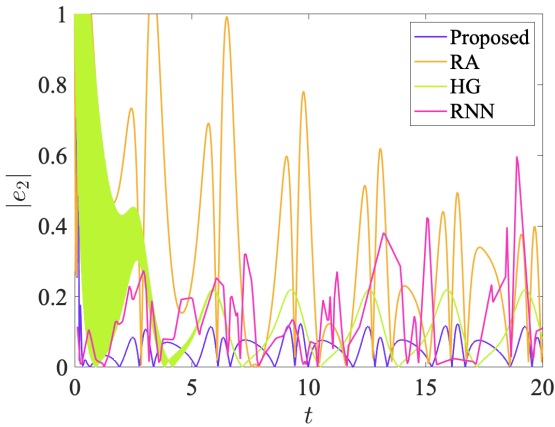


Figure 3: Comparative study of our proposed method with robust and adaptive observers in the literature. RA = Robust adaptive observer [28, Chapter 5], HG = High-gain observer [29], RNN = Recurrent neural network observer [30]. RA and HG have knowledge of the nonlinearity ϕ , and RNN state is initialized near the system’s initial state.

7. Conclusions

In this paper, we provide a design methodology for constructing state estimators with unmodeled dynamics. We generate an initial observer that is robust to the learning error, in the sense of ISS guarantees, and use this conservative initial state estimator to iteratively learn the unmodeled dynamics via Bayesian optimization in a purely

data-driven manner. Once a suitable estimate of the complete model is achieved, based on an information measure, we redesign the observer using the learned dynamics to reduce conservatism. Furthermore, the proposed modular design, based on a model-based part ensuring the safety of the learning via ISS guarantees, added to a data-driven optimization part, allows us to plug-in any data-driven optimization approach with convergence guarantees, e.g. Bayesian optimization, extremum seekers, etc.

Further investigations will be conducted to deal with the case of output measurement noise, which can enter the system both at the observer feedback level or at the learning cost function measurement level.

References

- [1] K. S. Narendra, K. Parthasarathy, Identification and control of dynamical systems using neural networks, *IEEE Transactions on Neural Networks* 1 (1990) 4–27.
- [2] E. B. Kosmatopoulos, M. M. Polycarpou, M. A. Christodoulou, P. A. Ioannou, High-order neural network structures for identification of dynamical systems, *IEEE Transactions on Neural Networks* 6 (1995) 422–431.
- [3] A. Chakrabarty, V. Dinh, M. J. Corless, A. E. Rundell, S. H. Žak, G. T. Buzzard, Support vector machine informed explicit nonlinear model predictive control using low-discrepancy sequences, *IEEE Transactions on Automatic Control* 62 (2017) 135–148.
- [4] J. Drgoňa, D. Picard, M. Kvasnica, L. Helsen, Approximate model predictive building control via machine learning, *Applied Energy* 218 (2018) 199–216.
- [5] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, F. L. Lewis, Optimal and autonomous control using reinforcement learning: A survey, *IEEE Transactions on Neural Networks and Learning Systems* 29 (2018) 2042–2062.
- [6] Y. Huang, D. Liu, Neural-network-based optimal tracking control scheme for a class of unknown discrete-time nonlinear systems using iterative adp algorithm, *Neurocomputing* 125 (2014) 46–56.
- [7] J. Theocharis, V. Petridis, Neural network observer for induction motor control, *IEEE Control Systems Magazine* 14 (1994) 26–37.
- [8] Y. H. Kim, F. L. Lewis, C. T. Abdallah, A dynamic recurrent neural-network-based adaptive observer for a class of nonlinear systems, *Automatica* 33 (1997) 1539–1543. doi:10.1016/S0005-1098(97)00065-4.
- [9] F. Abdollahi, H. Talebi, R. Patel, A Stable Neural Network-Based Observer With Application to Flexible-Joint Manipulators, *IEEE Transactions on Neural Networks* 17 (2006) 118–129. doi:10.1109/TNN.2005.863458.
- [10] X. Liu, C. Yang, Z. Chen, M. Wang, C. Y. Su, Neuro-adaptive observer based control of flexible joint robot, *Neurocomputing* 275 (2018) 73–82. doi:10.1016/j.neucom.2017.05.011.
- [11] B. Pratap, S. Purwar, Neural Network Observer for Twin Rotor MIMO System: An LMI Based Approach, in: *International Conference on Modeling, Identification and Control*, IEEE, Okayama, Japan, 2010, pp. 539–544.
- [12] R. Rahimilarki, Z. Gao, A. Zhang, R. J. Binns, Robust neural network fault estimation approach for nonlinear dynamic systems with applications to wind turbine systems, *IEEE Transactions on Industrial Informatics* 3203 (2019) 1–1.
- [13] M. Benosman, J. Borggaard, Robust nonlinear state estimation for a class of infinite-dimensional systems using reduced-order models, *International Journal of Control* (2019) 1–12.
- [14] S. Koga, M. Benosman, J. Borggaard, Learning-based robust observer design for coupled thermal and fluid systems, in: *2019 American Control Conference (ACC)*, IEEE, 2019, pp. 941–946.

- [15] A. Chakrabarty, A. Zemouche, R. Rajamani, M. Benosman, Robust data-driven neuro-adaptive observers with lipschitz activation functions, in: 2019 IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 2862–2867.
- [16] M. Benosman, Model-based vs. data-driven adaptive control: An overview, *Journal of Adaptive Control and Signal Processing* 32 (2018) 753–776.
- [17] J. Na, G. Herrmann, K. G. Vamvoudakis, Adaptive optimal observer design via approximate dynamic programming, in: 2017 American Control Conference (ACC), IEEE, 2017, pp. 3288–3293.
- [18] A. Capone, S. Hirche, Interval observers for a class of nonlinear systems using gaussian process models, in: 2019 18th European Control Conference (ECC), 2019, pp. 1350–1355.
- [19] Y. N. Raffoul, Y. M. Dib, Boundedness and Stability in Non-linear Discrete Systems With Nonlinear Perturbation, *Journal of Difference Equations and Applications* 9 (2003) 853–862. doi:10.1080/1023619021000047699.
- [20] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Mathematics of control, signals and systems* 2 (1989) 303–314.
- [21] B. Hanin, Universal function approximation by deep neural nets with bounded width and relu activations, *Mathematics* 7 (2019) 992.
- [22] H. Li, L. Grüne, Computation of local ISS Lyapunov functions for discrete-time systems via linear programming, *Journal of Mathematical Analysis and Applications* 438 (2016) 701–719.
- [23] A. Mironchenko, Local input-to-state stability: Characterizations and counterexamples, *Syst. Control Lett.* 87 (2016) 23–28.
- [24] C. K. Williams, C. E. Rasmussen, *Gaussian Processes For Machine Learning*, MIT press. Cambridge, MA, 2006.
- [25] V. Nguyen, S. Gupta, S. Rana, C. Li, S. Venkatesh, Regret for expected improvement over the best-observed value and stopping condition, *Journal of Machine Learning Research* 77 (2017) 279–294.
- [26] V. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, 1976.
- [27] A. Chakrabarty, D. K. Jha, G. T. Buzzard, Y. Wang, K. G. Vamvoudakis, Safe approximate dynamic programming via kernelized Lipschitz estimation, *IEEE Transactions on Neural Networks and Learning Systems* (2020) 1–15. URL: DOI:10.1109/TNNLS.2020.2978805.
- [28] P. A. Ioannou, J. Sun, *Robust adaptive control*, Courier Corporation, 2012.
- [29] H. K. Khalil, L. Praly, High-gain observers in nonlinear feedback control, *International Journal of Robust and Nonlinear Control* 24 (2014) 993–1015.
- [30] B. Chen, H. Zhang, X. Liu, C. Lin, Neural observer and adaptive neural control design for a class of nonlinear systems, *IEEE Trans. on Neural Networks and Learning Systems* 29 (2017) 4261–4271.

Appendix: Implementation

Algorithm 1 Bayesian Learning-based Observer

- Require:** Initial conditions x_0, \hat{x}_0
Require: System matrices A, B, C
Require: Function argument matrix C_q \triangleright default: I
Require: Basis functions ψ
Require: Initial coefficients p_0 \triangleright default: 0
Require: Upper bound on coefficients \bar{p}_* \triangleright default: 0.01
Require: Range of $\hat{\mathcal{L}}_\psi$ to perform line search
Require: Batch size for learning T_ℓ
Require: Cost function weights W_1, W_2 \triangleright p.s.d. matrices
Require: Kernel \mathcal{K} for GP regression \triangleright default: Matérn 5/2
Require: Acquisition Function \triangleright default: EI
Require: Termination threshold ε_{EI} \triangleright default: 0.001
-

Initial L-ISS observer design

- 1: Select A such that (A, C) is observable
 - 2: $\hat{\mathcal{L}}_\psi, L_0 \leftarrow$ perform line search (14) \triangleright involves solving SDP (10) for L_0 corresponding to $\hat{\mathcal{L}}_\psi$
 - 3: Parameterize initial observer (2) with L_0 and p_0
-

Safe learning from online data

- 4: **for** $j = 0 : N - 1$ **do**
 - 5: Collect batch data (estimated states $\{\hat{x}\}$ and measurements $\{y\}$) from system
 - 6: $\mathcal{J}(p_j) \leftarrow$ compute via (15)
 - 7: Learn surrogate GP model from data $\{p_j, \mathcal{J}(p_j)\}$
 - 8: $\mu, \sigma \leftarrow$ use GP to compute on samples drawn from \mathbb{P}
 - 9: $p_{j+1} \leftarrow$ obtain via acquisition function (20)
 - 10: Re-parameterize observer (2) with L_0 and $p_t \leftarrow p_{j+1}$
 - 11: Terminate if condition (21) holds
 - 12: $p_\infty \leftarrow$ terminal value of p_{j+1}
-

Observer gain redesign

- 13: $L \leftarrow$ redesigned observer gain in (27)
 - 14: Parameterize re-designed observer (25) with L and p_∞
-