Concurrent learning for parameter estimation using dynamic state-derivative estimators

Rushikesh Kamalapurkar, Benjamin Reish, Girish Chowdhary, and Warren E. Dixon

***Abstract*—A concurrent learning (CL)-based parameter estimator is developed to identify the unknown parameters in a nonlinear system. Unlike state-of-the-art CL techniques that assume knowledge of the state-derivative or rely on numerical smoothing, CL is implemented using a dynamic state-derivative estimator. A novel purging algorithm is introduced to discard possibly erroneous data recorded during the transient phase for concurrent learning. Asymptotic convergence of the error states to the origin is established under a *persistent* excitation condition, and the error states are shown to be uniformly ultimately bounded (UUB) under a *finite* excitation condition.**

***Index Terms*—concurrent learning, parameter estimation, adaptive systems, observers, Lyapunov methods**

I. Introduction

Modeling and identification of input-output relationships of nonlinear dynamical systems has been a long-standing active area of research. A variety of offline techniques have been developed for system identification; however, when models are used for feedback control, the ability to adapt to changes in the environment and the ability to learn from input-output data are desirable. Motivated by applications in feedback control, online system identification techniques are investigated in results such as [1]–[4] and the references therein.

Parametric methods such as linear parameterization, neural networks and fuzzy logic systems approximate the system identification problem by a finite-dimensional parameter estimation problem, and hence, are popular tools for online nonlinear system identification. Parametric models have been widely employed for adaptive control of nonlinear systems. In general, adaptive control methods do not require or guarantee convergence of the parameter estimates to their true values. However, it has been shown that parameter convergence can improve robustness and transient performance of adaptive controllers (see, e.g., [5]–[8]). Parametric models have also been employed in optimal control techniques such as model-based predictive control (MPC) (see, e.g., [9]–[12]) and model-based reinforcement learning (MBRL) (see e.g., [13]–[16]). In MPC and MBRL, the controller is developed based on the parameter estimates; hence, stability of the closed-loop system and the performance of the developed controller critically depend on convergence of the parameter estimates to their ideal values.

Data-driven concurrent learning (CL) techniques are developed in results such as [8], [17], [18], where recorded data is concurrently used with online data to achieve parameter convergence under a relaxed *finite* excitation condition as opposed to the *persistent* excitation (PE) condition required by traditional adaptive control methods. CL techniques are motivated by the fact that a direct formulation of the parameter estimation error can be obtained provided the state-derivative is known or its estimate is otherwise available through techniques such as fixed-point smoothing [19]. The parameter estimation error can then be used in a gradient-based adaptation algorithm to drive the parameter estimates to their ideal values. If exact derivatives are not available, the parameter estimation error can be shown to decay to a neighborhood of the origin provided accurate estimates of the state-derivatives are available, where the size of the neighborhood depends on the derivative estimation error [19]. Experimental results such as [8] demonstrate that, since derivatives at past data points are required, noncausal numerical smoothing techniques can be used to generate satisfactory estimates of state-derivatives. Under Gaussian noise, smoothing is guaranteed to result in the best possible linear estimate corresponding to the available data [20, Section 5.3]; however, in general, the derivative estimation error resulting from numerical smoothing can not be quantified a priori. Furthermore, numerical smoothing requires additional processing and storage of data over a time-window that contains the point of interest. Hence, the problem of achieving parameter convergence under relaxed excitation conditions without using numerical differentiation is motivated.

In this technical note, an observer is employed to estimate the state-derivative. The derivative estimate generated by the observer converges exponentially to a neighborhood of the actual state-derivative. However, in the transient phase, the derivative estimation errors can be large. Since CL relies on repeated use of recorded data, large transient errors present a challenge in the development of a CL-based parameter estimator. If the derivative estimation errors at the points recorded in the history stack are large, then the corresponding errors in the parameter estimates will be large. Motivated by the results in [21] and [22], the aforementioned challenge is addressed in this technical note by designing a novel purging algorithm to purge possibly erroneous data from the history stack.

The PE condition can be shown to be sufficient to ensure that enough data can be recorded to populate the history stack after each purge. Since PE can be an impractical requirement in many applications, this technical note examines the behavior of the switched error system under a relaxed finite excitation condition. Specifically, provided the system states are exciting over a sufficiently long finite time-interval, the error states decay to an ultimate bound. Furthermore, the ultimate bound can be made arbitrarily small by increasing the learning gains. Simulation results are provided to demonstrate the effectiveness of the developed method under measurement noise.

Rushikesh Kamalapurkar and Benjamin Reish are with the Department of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK, USA. Email: {rushikesh.kamalapurkar, reish}@okstate.edu. Warren E. Dixon is with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA. Email: wdixon@ufl.edu. Girish Chowdhary is with the Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, IL, USA. Email: girishc@illinois.edu

This research is supported in part by NSF award number 1509516, ONR grant number N00014-13-1-0151, AFOSR Award Number FA9550-14-1-0399, and a contract with the AFRL, Munitions Directorate at Eglin AFB. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring agency.

II. Problem formulation

The system dynamics are assumed to be nonlinear and uncertain, described by the differential equation[1]

$$\dot{x} = f(x,u) = f^o(x,u) + g(x,u), \tag{1}$$

where the function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz continuous, $f^o : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents a portion of the dynamics which may be known, and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents unknown dynamics. On a compact set $\chi \subset \mathbb{R}^n \times \mathbb{R}^m$, the unknown function g can be approximated using basis functions as $g(x,u) = \theta^T \sigma(x,u) + \varepsilon(x,u)$, where $\sigma : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^P$ denotes the vector of basis functions, $\varepsilon : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ denotes the function approximation error, and $\theta \in \mathbb{R}^{P \times n}$ denotes the unknown constant parameter vector. The universal function approximation property of single layer neural networks can be used to conclude that, given a constant $\overline{\varepsilon}$, there exists a vector of basis functions, σ, a matrix of ideal parameters, θ, and positive constants $\overline{\sigma}$, and $\overline{\theta}$, such that $\|\theta\| < \overline{\theta}$, $\|\varepsilon(x,u)\| < \overline{\varepsilon}$, $\|\nabla_x \varepsilon(x,u)\| < \overline{\varepsilon}$, $\|\nabla_u \varepsilon(x,u)\| < \overline{\varepsilon}$, $\|\sigma(x,u)\| < \overline{\sigma}$, $\|\nabla_x \sigma(x,u)\| < \overline{\sigma}$, and $\|\nabla_u \sigma(x,u)\| < \overline{\sigma}$, $\forall (x,u) \in \chi$ [23].[2] The objective is to design a parameter estimator to estimate the unknown parameters. The system input is assumed to be a stabilizing controller such that $x, \dot{x}, u, \dot{u} \in \mathcal{L}_\infty$. The system state x is assumed to be available for feedback, and the state-derivative $\dot{x}$ is assumed to be unknown.

Let $\hat{x} \in \mathbb{R}^n$ and $\dot{\hat{x}} \in \mathbb{R}^n$ denote estimates of the measurable state, x, and the unmeasurable state-derivative, $\dot{x}$, respectively. Let $\hat{\theta} \in \mathbb{R}^{P \times n}$ denote an estimate of the unknown matrix θ. To achieve convergence of the estimate $\hat{\theta}$, to the ideal parameter matrix, θ, a CL-based parameter estimator is designed. The motivation behind CL is to adjust the parameters based on an estimate of the parameter identification error, $\tilde{\theta} \triangleq \theta - \hat{\theta}$, in addition to the state estimation error, $\tilde{x} \triangleq x - \hat{x}$. Since $\tilde{\theta}$ is not directly measurable, the subsequent development exploits the fact that the term $\tilde{\theta}^T \sigma(x,u)$ can be expressed as $\tilde{\theta}^T \sigma(x,u) = \dot{x} - f^o(x,u) - \hat{\theta}^T \sigma(x,u) - \varepsilon(x,u)$, provided measurements of the state-derivative are available. In CL results such as [8], [17], and [18], it is assumed that the state-derivatives can be computed with sufficient accuracy at a past time instance by numerically differentiating the recorded data. An approximation of the parameter estimation error is then expressed as $\tilde{\theta}^T \sigma(x_j,u_j) + \overline{d} + \varepsilon(x_j,u_j) = \dot{\overline{x}}_j - f^o(x_j,u_j) - \hat{\theta}^T \sigma(x_j,u_j)$, where x_j denotes the system state at a past time instance t_j, $\overline{x}_j$ denotes the numerically computed state-derivative at t_j, and $\overline{d}$ is a constant of the order of the error between $\dot{x}_j$ and $\dot{\overline{x}}_j$. While the results in [19] establish that, provided $\overline{d}$ is bounded, the parameter estimation error $\tilde{\theta}$ can be shown to decay to a ball around the origin, the focus is on the analysis of the effects of the differentiation error, and not on development of algorithms to reduce the parameter estimation error.

In this technical note, a dynamically generated estimate of the state-derivative is used instead of numerical smoothing. The parameter estimation error is computed at a past recorded data point as $\tilde{\theta}^T \sigma(x_j,u_j) - \dot{\tilde{x}}_j + \varepsilon(x_j,u_j) = \dot{\hat{x}}_j - f^o(x_j,u_j) - \hat{\theta}^T \sigma(x_j,u_j)$, where $\dot{\tilde{x}}_j \triangleq \dot{x}_j - \dot{\hat{x}}_j$. To facilitate the design, let $\mathcal{H} \triangleq \left\{ \left(\dot{\hat{x}}_j, x_j, u_j \right) \right\}_{j=1}^{M}$ be a history stack containing recorded values of the state, the control, and the state-derivative estimate. Each tuple $\left(\dot{\hat{x}}_j, x_j, u_j \right)$ is referred to as a data-point in $\mathcal{H}$. A history stack $\mathcal{H}$ is called "full rank" if the state vectors recorded in $\mathcal{H}$ satisfy rank $\left(\sum_{j=1}^{M} \sigma(x_j,u_j) \sigma^T(x_j,u_j) \right) = P$. Based on the subsequent Lyapunov-based stability analysis, the history stack is used to update the estimate $\hat{\theta}$ using the following update law:

$$\dot{\hat{\theta}} = \Gamma \sigma(x,u) \tilde{x}^T + k \Gamma \sum_{j=1}^{M} \sigma(x_j,u_j) \left(\dot{\hat{x}}_j - f^o(x_j,u_j) - \hat{\theta}^T \sigma(x_j,u_j) \right)^T, \tag{2}$$

where $\Gamma \in \mathbb{R}^{P \times P}$ and $k \in \mathbb{R}_{>0}$ are constant learning gains.[3] The matrix Γ is diagonal and positive definite, and can be used to adjust the learning rate for each individual parameter. The gain k can be used to adjust the contribution of concurrent learning.

The update law in (2) drives the parameter estimation error to a ball around the origin, the size of which, is of the order of $\dot{\tilde{x}}_j$. Hence, to achieve a lower parameter estimation error, it is desirable to drive $\dot{\tilde{x}}$ to the origin. Based on the Lyapunov-based stability analysis in Section IV, the following adaptive estimator is designed to generate the state-derivative estimates:[4]

$$\begin{aligned} \dot{\hat{x}} &= \gamma \hat{\theta}^T \sigma(x,u) + f^o(x,u) + (k_x + \alpha) \tilde{x} + \mu, \\ \dot{\mu} &= (k_x \alpha + 1) \tilde{x}, \end{aligned} \tag{3}$$

where $\mu \in \mathbb{R}^n$ is an auxiliary signal and $k_x, \alpha \in \mathbb{R}_{>0}$ and $\gamma \in [0,1]$ are positive constant learning gains.

III. Algorithm to record the history stack

A. Purging of history stacks

The state-derivative estimator in (3) relies on feedback of the state estimation error, $\tilde{x}$. In general, feedback results in large transient estimation errors. Hence, the state-derivative estimation errors associated with the tuples $\left(\dot{\hat{x}}_j, x_j, u_j \right)$ recorded in the transient phase can be large. The results in [19] indicate that the parameter estimation errors can be of the order of $\max_j \left\| \dot{\tilde{x}}_j \right\|$. Hence, if a history stack containing data-points with large derivative estimation errors is used for CL, then the parameter estimates converge but the resulting parameter estimation errors can be large. To address the aforementioned challenge, this technical note introduces a new algorithm that purges the erroneous data in the history stack as soon as more data is available. Since the estimator in (3) results in exponential convergence of $\dot{\hat{x}}$ to a neighborhood of $\dot{x}$, newer data is guaranteed to represent the system better than older data, resulting in a lower steady-state parameter estimation error. The following section details the proposed algorithm.

[1] Unless otherwise specified, an equation of the form $h = g(x,y,t) + f$ is interpreted as $h(t) = g(x(t), y(t), t) + f(t)$, $\forall t \in [0,\infty)$, and a definition of the form $h(x,t) \triangleq g(x,t) + f(x)$ is interpreted as $h(x,t) = g(x,t) + f(x)$, $\forall (x,t) \in \mathbb{R}^n \times [0,\infty)$.

[2] The notation $\|\cdot\|$ denotes the Euclidean norm for vectors and the Frobenius norm for matrices. The notation $\nabla_{(\cdot)}$ represents the partial derivative $\frac{\partial}{\partial (\cdot)}$.

[3] For $a \geq 0$, $\mathbb{R}_{>a}$ and $\mathbb{R}_{\geq a}$ denote the intervals (a,∞) and $[a,\infty)$, respectively.

[4] A high gain state-derivative estimator is used in this technical note for ease of exposition. High order sliding mode differentiators (see e.g., [24]–[27]) could potentially be utilized to estimate $\dot{x}$ more efficiently.

B. Algorithm to record the history stack

The history stack $\mathcal{H}$ is initialized arbitrarily to be full rank. An arbitrary full rank initialization of $\mathcal{H}$ results in a $\sigma-$modification (cf. [28]) like adaptive update law that keeps the parameter estimation errors bounded. The data collected from the system is recorded in an auxiliary history stack $\mathcal{G} \triangleq \left\{\left(\dot{\hat{x}}_j^G, x_j^G, u_j^G\right)\right\}_{j=1}^{M}$. The history stack $\mathcal{G}$ is initialized such that $\left(\dot{\hat{x}}_j^G(0), x_j^G(0), u_j^G(0)\right) = (0,0,0)$ and is populated using a singular value maximization algorithm [17]. When all the elements of the history stack $\mathcal{G}$ have been replaced at least once and it becomes full rank with a minimum singular value that is above a (static or dynamic) threshold, $\mathcal{H}$ is replaced with $\mathcal{G}$, and $\mathcal{G}$ is purged.[5] In this technical note, a dynamic threshold is used which is set to be a fraction of the highest encountered minimum singular value corresponding to $\mathcal{H}$ up to the current time.

In the subsequent Algorithm 1, a piece-wise constant function $\delta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, initialized to zero, stores the last time instance when $\mathcal{H}$ was updated and a piecewise constant function $\eta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ stores the highest encountered value of $s_{\min}\left(\sum_{j=1}^{M} \sigma(x_j, u_j)\sigma^T(x_j, u_j)\right)$ up to time t, where $s_{\min}$ denotes the minimum singular value. The constant $\xi \in (0,1)$ denotes the threshold fraction used to purge the history stack, and $\mathcal{T} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a piecewise constant function.

Algorithm 1 History stack purging with dwell time

if a data point is available **then**
  if $\mathcal{G}$ is not full **then**
    add the data to $\mathcal{G}$
  else
    add the data to $\mathcal{G}$ if $s_{\min}\left(\sum_{j=1}^{M} \sigma\left(x_j^G, u_j^G\right)\sigma^T\left(x_j^G, u_j^G\right)\right)$ increases
  end if
  if $s_{\min}\left(\sum_{j=1}^{M} \sigma\left(x_j^G, u_j^G\right)\sigma^T\left(x_j^G, u_j^G\right)\right) \geq \xi\eta(t)$ **then**
    if $t - \delta(t) \geq \mathcal{T}(t)$ **then**
      $\mathcal{H} \leftarrow \mathcal{G}$ and $\mathcal{G} \leftarrow 0$ (purge $\mathcal{G}$)
      $\delta(t) \leftarrow t$
      if $\eta(t) < s_{\min}\left(\sum_{j=1}^{M} \sigma(x_j, u_j)\sigma^T(x_j, u_j)\right)$ **then**
        $\eta(t) \leftarrow s_{\min}\left(\sum_{j=1}^{M} \sigma(x_j, u_j)\sigma^T(x_j, u_j)\right)$
      end if
    end if
  end if
end if

[5]Techniques such as probabilistic confidence checks [22] or exponentially decaying singular value bounds [21] can also be utilized to initiate purging. The following analysis is agnostic with respect to the trigger used for purging provided $\mathcal{G}$ is full rank at the time of purging and the dwell time $\mathcal{T}$ is maintained between two successive purges.

IV. Analysis

Parameter convergence is typically established assuming that the system states are persistently exciting. However, from a practical perspective, excitation beyond a certain finite time-interval may not be available. If excitation is available only over a finite time-interval, then the parameter estimation errors can be made as small as desired, provided the history stacks are updated so that the time interval between two consecutive updates, i.e., the dwell-time, is large enough according to the conditions developed in Theorem 1.

A. Ultimate boundedness under finite excitation

The updates $\mathcal{H} \leftarrow \mathcal{G}$ imply that the resulting closed-loop system is a switched system, where each subsystem corresponds to a history stack, and each update indicates a switching event.[6] To facilitate the analysis, let $\rho : \mathbb{R}_{\geq 0} \to \mathbb{N}$ denote a switching signal such that $\rho(0) = 1$, and $\rho(t) = i+1$, where i denotes the number of times the update $\mathcal{H} \leftarrow \mathcal{G}$ was carried out over the time interval $(0, t)$. In the following, the subscript $s \in \mathbb{N}$ denotes the switching index, and $\mathcal{H}_s$ denotes the history stack corresponding to the s^{th} subsystem (i.e., the history stack active during the time interval $\{t \mid \rho(t) = s\}$), containing the elements $\left\{\left(\dot{\hat{x}}_{sj}, x_{sj}, u_{sj}\right)\right\}_{j=1}^{M}$. To simplify the notation, let

$$A_s = \sum_{j=1}^{M} \sigma(x_{sj}, u_{sj})\sigma^T(x_{sj}, u_{sj}),$$

$$Q_s = \sum_{j=1}^{M} \sigma(x_{sj}, u_{sj})\left(\dot{\hat{x}}_{sj}^T + \varepsilon^T(x_{sj}, u_{sj})\right).$$

Note that $A_s : \mathbb{R}_{\geq 0} \to \mathbb{R}^{P\times P}$ and $Q_s : \mathbb{R}_{\geq 0} \to \mathbb{R}^{P\times 1}$ are piece-wise constant functions of time.

Algorithm 1 ensures that there exists a constant $\underline{a} > 0$ such that $\lambda_{\min}\{A_s\} \geq \underline{a}$, $\forall s \in \mathbb{N}$, where $\lambda_{\min}$ denotes the minimum eigenvalue. Since the state x remains bounded by assumption, there exists a constant $\overline{A}$ such that $\|A_s\| \leq \overline{A}$, $\forall s \in \mathbb{N}$.

Using (2), the dynamics of the parameter estimation error, $\tilde{\theta}$, can be written as

$$\dot{\tilde{\theta}} = -\Gamma\sigma(x,u)\tilde{x}^T - k\Gamma A_s\tilde{\theta} + k\Gamma Q_s. \tag{4}$$

To establish convergence of the state-derivative estimates, a filtered tracking error, $r \in \mathbb{R}^n$, is defined as $r \triangleq \dot{\tilde{x}} + \alpha\tilde{x}$. Using (1), (3), and (4), the time derivative of the filtered tracking error can be expressed as,

$$\dot{r} = -\gamma\tilde{x}\sigma^T(x,u)\Gamma\sigma(x,u) - k\gamma\tilde{\theta}^T A_s^T\Gamma\sigma(x,u) - k_x r - \tilde{x} + k\gamma Q_s^T\Gamma\sigma(x,u) + \gamma\tilde{\theta}^T F(x,u,\dot{u}) + E(x,u,\dot{u}) + (1-\gamma)\theta^T F(x,u,\dot{u}), \tag{5}$$

where $F(x,u,\dot{u}) \triangleq \nabla_x\sigma(x,u)f(x,u) + \nabla_u\sigma(x,u)\dot{u}$ and $E(x,u,\dot{u}) \triangleq \nabla_x\varepsilon(x,u)f(x,u) + \nabla_u\varepsilon(x,u)\dot{u}$. To facilitate

[6]Since a switching event in Algorithm 1 occurs only when the auxiliary history stack is full, Zeno behavior is avoided by design.

the stability analysis, let $\overline{F}$, $\overline{F}^o$, $\overline{x}$, and $\overline{\Gamma}$ be constants such that

$$\|F(x,u,\dot{u})\| \leq \overline{F},\ \|\Gamma\| = \overline{\Gamma},$$
$$\left\|f^o(x,u) + \theta^T \sigma(x,u) + \varepsilon(x,u)\right\| \leq \overline{F}^o,\ \|x\| \leq \overline{x}, \quad (6)$$

for all $t \in \mathbb{R}_{\geq 0}$. The following stability analysis is split into three parts. In Part 1, it is established that provided the error states $\tilde{\theta}$, $\tilde{x}$, and r are bounded at a switching instance and that the norms of the state-derivative estimates stored in the history stack are bounded, $\tilde{\theta}$, $\tilde{x}$, and r decay to a bound before the next switching instance, where the bound depends on the derivative estimation errors. In Part 2, it is established that provided the error states $\tilde{x}$ and r are bounded at a switching instance, the derivative estimation error $\dot{\tilde{x}}$ can be made arbitrarily small before the next switching instance by increasing the learning gains. In Part 3, the results of Part 1 and Part 2 are used to conclude convergence of the error state $\tilde{\theta}$ to a ball around the origin.

Part 1: Boundedness of the error signals: Let $Z \triangleq \begin{bmatrix} r^T & \tilde{x}^T & \text{vec}\left(\tilde{\theta}\right)^T \end{bmatrix}^T \in \mathbb{R}^{2n+L}$ and let $V : \mathbb{R}^{2n+P} \to \mathbb{R}_{\geq 0}$ denote a candidate Lyapunov function defined as

$$V(Z) \triangleq \frac{1}{2} r^T r + \frac{1}{2}\tilde{x}^T\tilde{x} + \frac{1}{2}\text{tr}\left(\tilde{\theta}^T\Gamma^{-1}\tilde{\theta}\right), \quad (7)$$

where $\text{vec}(\cdot)$ denotes the vectorization operator. Using the Raleigh-Ritz Theorem and the fact that $\text{tr}\left(\tilde{\theta}^T\Gamma^{-1}\tilde{\theta}\right) = \left(\text{vec}\left(\tilde{\theta}\right)\right)^T\left(\Gamma^{-1}\otimes\mathbb{I}_P\right)\left(\text{vec}\left(\tilde{\theta}\right)\right)$, the Lyapunov function V can be bounded as

$$\underline{v}\|Z\|^2 \leq V(Z) \leq \overline{v}\|Z\|^2, \quad (8)$$

where $\overline{v} \triangleq \frac{1}{2}\max\left\{1, \lambda_{\max}\left\{\Gamma^{-1}\right\}\right\}$, $\underline{v} \triangleq \frac{1}{2}\min\left\{1, \lambda_{\min}\left\{\Gamma^{-1}\right\}\right\}$. The subsequent stability analysis assumes that the learning gains k, k_x, and α, and the matrices A_s satisfy the following sufficient gain conditions:[7]

$$\underline{a} > \frac{3\overline{\sigma}^2}{k\alpha} + \frac{4\gamma^2\overline{F}^2}{kk_x} + \frac{4k\gamma^2\overline{\sigma}^2\overline{\Gamma}^2\overline{A}^2}{k_x}, \quad k_x > \frac{6\gamma^2\overline{\sigma}^4\overline{\Gamma}^2}{\alpha}. \quad (9)$$

The following lemma focuses on a time interval $[\tau, \tau+T)$. The objective of the lemma is to find an envelope that bounds the trajectory of the candidate Lyapunov function over the interval.

Lemma 1. *Let τ, $T \in \mathbb{R}_{\geq 0}$ be constants such that $\rho(t) = s$, for all $t \in [\tau, \tau+T)$. Let $\overline{V}_s \in \mathbb{R}_{>0}$ be a constant such that the candidate Lyapunov function V satisfies $V(Z(\tau)) \leq \overline{V}_s$. Then, V is bounded over the entire interval $[\tau, \tau+T)$ as*

$$V(Z(t)) \leq \left(\overline{V}_s - \frac{\overline{v}}{v}\iota_s\right)e^{-\frac{v}{\overline{v}}(t-\tau)} + \frac{\overline{v}}{v}\iota_s,\ \forall t \in [\tau, \tau+T), \quad (10)$$

where $\iota_s \triangleq \left(\frac{k}{2\underline{a}} + \frac{k^2\gamma^2\overline{\sigma}^2\overline{\Gamma}^2}{k_x}\right)\overline{Q}_s^2 + \frac{2\left(\overline{F}\overline{\theta}+\overline{E}\right)^2(1-\gamma)^2}{k_x}$, $v \triangleq \min\left\{\frac{k\underline{a}}{4}, \frac{\alpha}{3}, \frac{k_x}{8}\right\}$, and $\overline{Q}_s \triangleq \|Q_s\|$. Furthermore, the parameter estimation error can be bounded as

$$\left\|\tilde{\theta}(t)\right\| \leq \theta_s,\ \forall t \in [\tau, \tau+T) \quad (11)$$

where $\theta_s \triangleq \sqrt{\frac{\overline{v}}{\underline{v}}}\max\left\{\sqrt{\overline{V}_s}, \sqrt{\iota_s}\right\}$.

Proof: Provided the sufficient conditions in (9) are satisfied, (4) and (5) can be used to bound the time derivative of the candidate Lyapunov function V as

$$\dot{V} \leq -\frac{v}{\overline{v}}V(Z) + \iota_s.$$

Using the comparison lemma [29, Lemma 3.4],

$$V(Z(\tau)) \leq \left(\overline{V}_s - \frac{\overline{v}}{v}\iota_s\right)e^{-\frac{v}{\overline{v}}(t-\tau)} + \frac{\overline{v}}{v}\iota_s,\ \forall t \in [\tau, \tau+T).$$

If $\overline{V}_s \geq \frac{\overline{v}}{v}\iota_s$ then $V(Z(\tau)) \leq \overline{V}_s$. If $\overline{V}_s < \frac{\overline{v}}{v}\iota_s$ then $V(Z(\tau)) \leq \frac{\overline{v}}{v}\iota_s$. Hence, using the definition of Z and the bounds in (8), the bound in (11) is obtained. ∎

Part 2: Exponential decay of $\dot{\tilde{x}}$: Let $Z_r \triangleq \begin{bmatrix} r^T & \tilde{x}^T \end{bmatrix}^T \in \mathbb{R}^{2n}$ and let $V_r : \mathbb{R}^{2n} \to \mathbb{R}_{\geq 0}$ be a candidate Lyapunov function defined as $V_r(Z_r) \triangleq \|Z_r\|^2$. The following lemma establishes exponential convergence of the derivative estimation error to a neighborhood of the origin.

Lemma 2. *Let all the conditions of Lemma 1 be satisfied. Let $\overline{V}_{rs}$ be a constant such that the candidate Lyapunov function V_r satisfies $V_r(Z_r(\tau)) \leq \overline{V}_{rs}$. Then, V_r is bounded over the entire interval $[\tau, \tau+T)$ as*

$$V_r(Z_r(t)) \leq \left(\overline{V}_{rs} - \frac{\iota_{rs}}{v_r}\right)e^{-v_r(t-\tau)} + \frac{\iota_{rs}}{v_r},\ \forall t \in [\tau, \tau+T), \quad (12)$$

where $v_r = \min\left\{\frac{k_x}{2}, \alpha\right\}$ and $\iota_{rs} \triangleq \frac{4\left(\overline{E}+(1-\gamma)\overline{\theta F}\right)^2}{k_x} + \frac{4\gamma^2\left(k\overline{A\Gamma}\overline{\sigma}+\overline{F}\right)^2}{k_x}\overline{\theta}_s^2 + \frac{4\gamma^2k^2\overline{\Gamma}^2\overline{\sigma}^2}{k_x}\overline{Q}_s^2$. Furthermore, given a constant $\epsilon_r \in \mathbb{R}_{>0}$, the gain k_x can be selected large enough such that $\left\|\dot{\tilde{x}}(\tau+T)\right\| \leq \epsilon_r$.

Proof: Provided the sufficient conditions in (9) are satisfied, (5) and Lemma 1 can be used to bound the time-derivative of the candidate Lyapunov function V_r as

$$\dot{V}_r \leq -v_r V_r(Z_r) + \iota_{rs}.$$

Using the comparison lemma, [29, Lemma 3.4]

$$V_r(Z_r(t)) \leq \left(\overline{V}_{rs} - \frac{\iota_{rs}}{v_r}\right)e^{-v_r(t-\tau)} + \frac{\iota_{rs}}{v_r},\ \forall t \in [\tau, \tau+T).$$

Using the fact that $\dot{\tilde{x}} = r - \alpha\tilde{x}$, the state-derivative estimation error can be bounded as $\left\|\dot{\tilde{x}}\right\|^2 \leq \|r\|^2 + \alpha\|\tilde{x}\|^2 \leq (1+\alpha)V_r(Z_r)$. Based on (12), given $\overline{V}_{rs} \geq V_r(Z_r(t))$, $\epsilon_r > 0$, the gain k_x can be selected large enough so that $V_r(Z_r(\tau+T)) \leq \frac{\epsilon_r^2}{(1+\alpha)}$. Hence, given $\overline{V}_{rs}$, $\epsilon_r > 0$, the gain k_x can be selected to be large enough so that $\left\|\dot{\tilde{x}}(\tau+T)\right\| \leq \epsilon_r$. ∎

In the following, the results of Lemmas 1 and 2 are used in an inductive argument to show that all the states of the dynamical system defined by (4) and (5) remain bounded and decay to a ball around the origin provided enough data can be recorded to repopulate the history stack after at-least two purges.

[7] The sufficient conditions can be satisfied provided the gains k_x and α are selected large enough.

Part 3: Ultimate boundedness under finite excitation: To facilitate the analysis, let $\{T_s \in \mathbb{R}_{\geq 0} \mid s \in \mathbb{N}\}$ be a set of switching time instances defined as $T_s = \{t \mid \rho(\tau) < s+1,\ \forall \tau \in [0,t) \wedge \rho(\tau) \geq s+1,\ \forall \tau \in [t,\infty)\}$. That is, for a given switching index s, T_s denotes the time instance when the $(s+1)^{\text{th}}$ subsystem is switched on. For notational brevity, let $\zeta_s \triangleq 4\left(\overline{E}+(1-\gamma)\overline{\theta F}\right)^2 + 4\gamma^2\left(k\overline{\sigma\Gamma}\right)^2\overline{Q}_s^2 + 4\gamma^2\left(k\overline{A\Gamma}\overline{\sigma}+\overline{F}\right)^2\overline{\theta}_s^2$, $\varpi \triangleq 2\left(\overline{F\theta}+\overline{E}\right)^2(1-\gamma)^2$, and $\beta \triangleq \left(\frac{k}{2\underline{a}} + \frac{k^2\gamma^2\overline{\sigma}^2\overline{\Gamma}^2}{k_x}\right)$.

Theorem 1. *Let $\epsilon > 0$ be given. Let the history stacks $\mathcal{H}$ and $\mathcal{G}$ be populated using Algorithm 1. Let the learning gains be selected to satisfy the sufficient gain conditions in (9) and the sufficient gain condition*

$$k_x > \max\left\{\frac{32 v_r \overline{v}^2 M\overline{\sigma}(1+\alpha)\zeta_1\beta}{\underline{v}^2 v^2 \epsilon^2}, \frac{4\overline{v}\varpi}{v\underline{v}\epsilon}\right\}. \tag{13}$$

Let $T \in \mathbb{R}_{>0}$ be a time instance such that the system states are exciting over $[0,T]$, that is, the history stack can be replenished if purged at any time $t \in [0,T]$. Assume that over each switching interval $\{t \mid \rho(t) = s\}$, the dwell-time, $\mathcal{T}$, is selected such that $\mathcal{T}(t) = \mathcal{T}_s$, where $\mathcal{T}_s$ is selected to be large enough to satisfy $\left(\overline{V}_s - \frac{\overline{v}}{v}\iota_s\right)e^{-\frac{v}{\overline{v}}(\mathcal{T}_s)} \leq \frac{\overline{v}}{v}\iota_s$. Furthermore assume that the excitation interval is large enough so that $T_s < T$ for some $s > 2$.[8] *Then, $\limsup_{t\to\infty}\left\|\tilde{\theta}(t)\right\| \leq \epsilon$.*

Proof: Since the history stack $\mathcal{H}_1$ is selected at random to contain bounded elements, Q_1 is bounded and all the conditions of Lemmas 1 and 2 are satisfied over the time interval $[0, T_1)$. Hence, $V(Z(T_1)) \leq \left(\overline{V}_1 - \frac{\overline{v}}{v}\iota_1\right)e^{-\frac{v}{\overline{v}}(T_1)} + \frac{\overline{v}}{v}\iota_1$, where $\iota_1 = \beta\overline{Q}_1^2 + \frac{\varpi}{k_x}$. Using the bounds in (6), $\overline{Q}_1$ can be computed as $\overline{Q}_1 = qM\overline{\sigma}\left(\overline{F}^o + \overline{H}_1\right)$, where $q > 1$ is an adjustable parameter. Furthermore, $V_r(Z_r(T_1)) \leq \left(\overline{V}_{r1} - \frac{\iota_{r1}}{v_r}\right)e^{-v_r(T_1)} + \frac{\iota_{r1}}{v_r}$, where $\iota_{r1} = \frac{\zeta_1}{k_x}$ and $\theta_1 = \sqrt{\frac{\overline{v}}{\underline{v}}}\max\left\{\sqrt{\overline{V}_1}, \sqrt{\iota_1}\right\}$. Provided T_1 is large enough so that $\left(\overline{V}_{r1} - \frac{\iota_{r1}}{v_r}\right)e^{-v_r(T_1)} \leq \frac{\iota_{r1}}{v_r}$ and $\left(\overline{V}_1 - \frac{\overline{v}}{v}\iota_1\right)e^{-\frac{v}{\overline{v}}(T_1)} \leq \frac{\overline{v}}{v}\iota_1$, then $V_r(Z_r(T_1)) \leq \frac{2\iota_{r1}}{v_r}$ and $V(Z(T_1)) \leq 2\frac{\overline{v}}{v}\iota_1$, and hence, Q_2 is bounded.

Let $\epsilon^o > 0$ be a constant, to be selected later. Provided the gain k_x is selected such that $k_x > \max\left\{\frac{2(1+\alpha)\zeta_1}{v_r\epsilon^{o2}}, \frac{4\overline{v}\varpi}{v\underline{v}\epsilon}\right\}$, then $\overline{V}_{r2} \leq \frac{\epsilon^{o2}}{(1+\alpha)}$. Since $T_2 < T$ by assumption, the Lyapunov envelope W_r decays over the time interval $[T_1, T_2)$; hence, $\sup_{t\in[T_1,T_2)}\left\|\dot{\tilde{x}}(t)\right\|^2 \leq (1+\alpha)\overline{V}_{r2}$, that is, $\sup_{t\in[T_1,T_2)}\left\|\dot{\tilde{x}}(t)\right\| \leq \epsilon^o$. Hence, Q_3 is bounded and the bound $\overline{Q}_3$ can be selected as $\overline{Q}_3 = M\overline{\sigma}\epsilon^o$, which implies $\iota_3 = \beta M\overline{\sigma}\epsilon^o + \frac{\varpi}{k_x}$. Selecting $\epsilon^o = \frac{v\underline{v}\epsilon}{4\overline{v}\beta M\overline{\sigma}}$, the inequality $\iota_3 \leq \frac{v\underline{v}\epsilon}{2\overline{v}}$ is obtained.

If the history stack is only purged twice, then over the interval $[T_2, \infty)$, the Lyapunov function satisfies $V(Z(t)) \leq \left(\overline{V}_3 - \frac{\overline{v}}{v}\iota_3\right)e^{-\frac{v}{\overline{v}}(t-T_2)} + \frac{\overline{v}}{v}\iota_3$. Hence, $\limsup_{t\to\infty} V(Z(t)) \leq \frac{\overline{v}}{v}\iota_3 = \frac{v\epsilon}{2}$. Hence, using the bounds on the Lyapunov function in (8), $\limsup_{t\to\infty}\left\|\tilde{\theta}(t)\right\| \leq \epsilon$.

The rest of the analysis concerns the case where the history stack is continually purged whenever enough new data to satisfy the rank condition is available. The challenge is to find conditions to avoid error growth during the continual update process. To begin with, the time-interval $[T_1, T_2)$ is considered. It is already established that $V(Z(T_1))$ is bounded by $\overline{V}_2 \triangleq 2\frac{\overline{v}}{v}\iota_1$. Hence, Lemma 1 can be applied over the time interval $[T_1, T_2)$ to conclude that $V(Z(T_2)) \leq \left(\overline{V}_2 - \frac{\overline{v}}{v}\iota_2\right)e^{-\frac{v}{\overline{v}}(T_2-T_1)} + \frac{\overline{v}}{v}\iota_2$. Provided $T_1 - T_2$ is large enough so that $\left(\overline{V}_2 - \frac{\overline{v}}{v}\iota_2\right)e^{-\frac{v}{\overline{v}}(T_2-T_1)} \leq \frac{\overline{v}}{v}\iota_2$, then $V(Z(T_2)) \leq 2\frac{\overline{v}}{v}\iota_2$, and hence, Q_3 is bounded. Similarly, Lemma 2 can be applied over the time interval $[T_1, T_2)$ to conclude that provided $T_1 - T_2$ is large enough so that $\left(\overline{V}_{r_2} - \frac{\iota_{r2}}{v_r}\right)e^{-v_r(T_2-T_1)} \leq \frac{\iota_{r2}}{v_r}$, then $V_r(Z_r(T_2)) \leq 2\frac{\iota_{r2}}{v}$. By selecting the constant q such that $q > \max\left\{1, \frac{\sqrt{2(1+\alpha)\overline{V}_{r1}}}{\left(\overline{F}^o + \overline{H}_1\right)}\right\}$, it can be ensured that $\overline{Q}_1 \geq \overline{Q}_2$, hence, $\iota_1 \geq \iota_2$ and $V(Z(T_1)) \geq V(Z(T_2))$. Hence, $\theta_1 \geq \theta_2$. The inequalities $\overline{Q}_1 \geq \overline{Q}_2$ and $\theta_1 \geq \theta_2$ imply that $\iota_{r1} \geq \iota_{r2}$, which in turn, implies that $V_r(Z_r(T_1)) \geq V_r(Z_r(T_2))$.

The inequality $V_r(Z_r(T_1)) \geq V_r(Z_r(T_2))$ and the equation $\overline{Q}_s = M\overline{\sigma}\sqrt{2(1+\alpha)\overline{V}_{r(s-1)}}$ imply that $\overline{Q}_2 \geq \overline{Q}_3$. Using Lemmas 1 and 2 along with the inequality $\overline{Q}_2 \geq \overline{Q}_3$, the inequalities relating the candidate Lyapunov functions can be propagated to any given interval $[T_{s-1}, T_s)$ such that $V_r(Z_r(T_1)) \geq V_r(Z_r(T_2)) \geq \cdots \geq V_r(Z_r(T_s))$, $V(Z(T_1)) \geq V(Z(T_2)) \geq \cdots \geq V(Z(T_s))$, $\iota_1 \geq \iota_2 \geq \cdots \geq \iota_s$, and $\iota_{r1} \geq \iota_{r2} \geq \cdots \geq \iota_{rs}$. In particular, for any given $s > 2$, over the interval $[T_{s-1}, \infty)$, the candidate Lyapunov function V satisfies $V(Z(t)) \leq \left(\overline{V}_s - \frac{\overline{v}}{v}\iota_s\right)e^{-\frac{v}{\overline{v}}(t-T_{s-1})} + \frac{\overline{v}}{v}\iota_s$. Hence, $\limsup_{t\to\infty} V(Z(t)) \leq \frac{\overline{v}}{v}\iota_s \leq \frac{\overline{v}}{v}\iota_3 = \frac{v\epsilon}{2}$. Hence, using (8), $\limsup_{t\to\infty}\left\|\tilde{\theta}(t)\right\| \leq \epsilon$. ∎

Remark 1. The proof of Theorem 1 requires that for any given s, the dwell time must be selected such that $T_{s-1} - T_s$ is large enough to satisfy $\left(\overline{V}_s - \frac{\overline{v}}{v}\iota_s\right)e^{-\frac{v}{\overline{v}}(T_s - T_{s-1})} \leq \frac{\overline{v}}{v}\iota_s$. That is, $\overline{V}_s e^{-\frac{v}{\overline{v}}(T_s-T_{s-1})} \leq \frac{\overline{v}}{v}\iota_s$. Since $\overline{V}_s$ is the bound on $V(Z(T_{s-1}))$ and ι_s depends on $\overline{Q}_s$, which in turn, depends on $\overline{V}_{s-1}$, all the information required to compute the dwell-time for the interval $[T_{s-1}, T_s)$ is available at at $t = T_{s-1}$.

B. Special case: asymptotic convergence

Theorem 1 establishes practical stability (i.e., uniform ultimate boundedness) of the error system. Asymptotic convergence of the error states to the origin can be established provided a perfect basis is available for the approximation of the function g, such that $\varepsilon(x,u) = 0$, for all $(x,u) \in \chi$, the system states are persistently excited such that the history stack $\mathcal{H}$ can always be replaced with a new, full rank history stack, and the constant γ, that multiplies the feed-forward part of the estimator, is set to 1.

Theorem 2. *Provided $\gamma = 1$, $\varepsilon(x,u) = 0$, for all $(x,u) \in \chi$, the history stacks $\mathcal{H}$ and $\mathcal{G}$ are populated using Algorithm 1,*

[8] A minimum of two purges are required to remove the randomly initialized data, and the data recorded during transient phase of the derivative estimator from the history stack.

a minimum dwell-time is maintained such that $\mathcal{T}(t) \geq \underline{\mathcal{T}}$, the sufficient gain conditions in (9) and (14) are satisfied, and the system states are exciting such that the history stack $\mathcal{H}$ can be persistently purged and replenished, i.e.,

$$s \to \infty, \text{ as } \quad t \to \infty,$$

then, $\left\|\tilde{\theta}(t)\right\| \to 0$, $\|r(t)\| \to 0$, *and* $\|\tilde{x}(t)\| \to 0$ *as* $t \to \infty$.

Proof: Since $\gamma = 1$, the constant ι_s can be expressed as $\iota_s = \frac{k}{2\underline{a}}\overline{Q}_s^2 + \frac{k^2\overline{\sigma}^2\overline{\Gamma}^2}{k_x}\overline{Q}_s^2$, i.e., $\iota_s = \frac{kM^2\overline{\sigma}^2(1+\alpha)^2\overline{V}_{r(s-1)}}{\underline{a}} + \frac{2k^2\overline{\sigma}^2\overline{\Gamma}^2M^2\overline{\sigma}^2(1+\alpha)^2\overline{V}_{r(s-1)}}{k_x}$. Let $\beta = \frac{kM^2\overline{\sigma}^2(1+\alpha)^2}{2\underline{a}} + \frac{2k^2\overline{\sigma}^2\overline{\Gamma}^2M^2\overline{\sigma}^2(1+\alpha)^2}{k_x}$. Then $\iota_s = \beta\overline{V}_{r(s-1)}$. Similarly, the constant ι_{rs} can be expressed as $\iota_{rs} = \frac{4\left(k\overline{A\Gamma}\overline{\sigma}+\overline{F}\right)^2}{k_x}\overline{\theta}_s^2 + \frac{4k^2\overline{\Gamma}^2\overline{\sigma}^2}{k_x}\overline{Q}_s^2$. That is, $\iota_{rs} = \frac{4\left(k\overline{\sigma\Gamma}\right)^2M^2\overline{\sigma}^2(1+\alpha)^2}{k_x}\overline{V}_{r(s-1)} + \frac{4\left(\overline{F}+k\overline{\sigma\Gamma A}\right)^2\overline{\theta}_s^2}{k_x}$. From Lemma 1, $\overline{\theta}_s^2 = \frac{\overline{v}}{\underline{v}}\max\left\{\overline{V}_s, \iota_s\right\} \leq \frac{\overline{v}}{\underline{v}}\overline{V}_s + \frac{\overline{v}}{\underline{v}}\beta\overline{V}_{r(s-1)}$. Using the definitions $\varpi \triangleq \frac{4\left(k\overline{\sigma\Gamma}\right)^2M^2\overline{\sigma}^2(1+\alpha)^2}{k_x} + \frac{4\overline{v}\beta\left(\overline{F}+k\overline{\sigma\Gamma A}\right)^2}{\underline{v}k_x}$ and $\zeta \triangleq \frac{4\overline{v}\left(\overline{F}+k\overline{\sigma\Gamma A}\right)^2\overline{\theta}_s^2}{\underline{v}k_x}$, the constant ι_{rs} can be bounded as $\iota_{rs} \leq \varpi\overline{V}_{r(s-1)} + \zeta\overline{V}_s$.

Lemmas 1 and 2, the bounds on ι_s and ι_{rs}, and the minimum dwell-time $\underline{\mathcal{T}}$ imply that the bounds $\overline{V}_s$ and $\overline{V}_{rs}$ on the candidate Lyapunov functions can be selected to satisfy the recurrence relations $\overline{V}_s = \overline{V}_{s-1}e^{-\frac{v}{\overline{v}}\underline{\mathcal{T}}} + \frac{\overline{v}}{v}\beta\overline{V}_{rs-2}$ and $\overline{V}_{rs} = \overline{V}_{r(s-1)}e^{-v_r\underline{\mathcal{T}}} + \frac{\varpi}{v_r}\overline{V}_{r(s-2)} + \frac{\zeta}{v_r}\overline{V}_{(s-1)}$. The recurrence relations form the discrete-time linear time-invariant system

$$\begin{bmatrix} \overline{V}_s \\ \overline{V}_{r(s-1)} \\ \overline{V}_{rs} \end{bmatrix} = \begin{bmatrix} e^{-\frac{v}{\overline{v}}\underline{\mathcal{T}}} & \frac{\overline{v}}{v}\beta & 0 \\ 0 & 0 & 1 \\ \frac{\zeta}{v_r} & \frac{\varpi}{v_r} & e^{-v_r\underline{\mathcal{T}}} \end{bmatrix} \begin{bmatrix} \overline{V}_{s-1} \\ \overline{V}_{r(s-2)} \\ \overline{V}_{r(s-1)} \end{bmatrix}.$$

Provided the dwell-time and the learning gains are selected to satisfy

$$\text{Re}\left\{\lambda_{\max}\begin{bmatrix} e^{-\frac{v}{\overline{v}}\underline{\mathcal{T}}} & \frac{\overline{v}}{v}\beta & 0 \\ 0 & 0 & 1 \\ \frac{\zeta}{v_r} & \frac{\varpi}{v_r} & e^{-v_r\underline{\mathcal{T}}} \end{bmatrix}\right\} < 1, \tag{14}$$

where $\text{Re}\{\cdot\}$ denotes the real part, then $\lim_{s\to\infty}\overline{V}_s = 0$ and $\lim_{s\to\infty}\overline{V}_{rs} = 0$. ∎

V. Simulation results

The developed technique is simulated using a model for a two-link robot manipulator arm. The uncertainty $g(x,u)$ is linearly parameterizable as $g^T(x,u) = \theta^T\sigma(x,u)$. That is, the selected model belongs to a sub-class of systems defined by (1), where the function approximation error, ε, is zero. Since the ideal parameters, θ, are known, the selected model facilitates quantitative analysis of the parameter estimation error. The four-dimensional state of the model is denoted by $x \triangleq \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T$. The dynamics of the model are described by (1), where

$$f^0(x,u) = \begin{bmatrix} x_3 \\ x_4 \\ -\left(M(x)\right)^{-1}V_m(x)\begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \left(M(x)\right)^{-1} \end{bmatrix} u,$$

TABLE I: Simulation results for the developed technique and numerical differentiation-based CL.

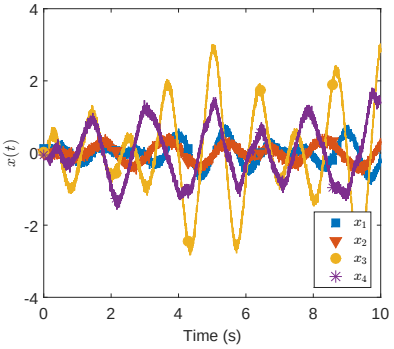
	Numerical differentiation-based CL		Developed technique	
Noise variance	0.005	0.1	0.005	0.1
RMS steady-state error	1.75	17.55	0.27	0.46

$$g^T(x,u) = \theta^T\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \left[\begin{array}{cc} \left(M(x)\right)^{-1} & \left(M(x)\right)^{-1} \end{array}\right] D(x) \end{bmatrix}^T. \tag{15}$$

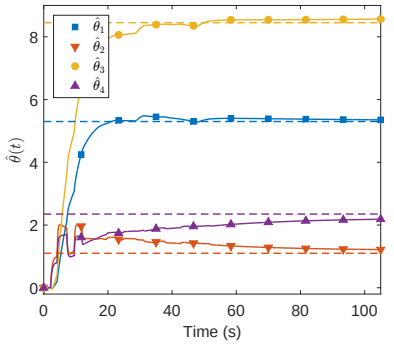
In (15), $u \in \mathbb{R}^2$ is the control input, $D(x) \triangleq \text{diag}\left[x_3,\, x_4,\, \tanh(x_3),\, \tanh(x_4)\right]$, $M(x) \triangleq \begin{bmatrix} p_1 + 2p_3c_2(x), & p_2 + p_3c_2(x) \\ p_2 + p_3c_2(x), & p_2 \end{bmatrix}$, and $V_m(x) \triangleq \begin{bmatrix} -p_3s_2(x)\,x_4, & -p_3s_2(x)\,(x_3 + x_4) \\ p_3s_2(x)\,x_3, & 0 \end{bmatrix}$, where $c_2(x) = \cos(x_2)$, $s_2(x) = \sin(x_2)$, and $p_1 = 3.473$, $p_2 = 0.196$, and $p_3 = 0.242$ are constants. The system has four unknown parameters. The ideal values of the unknown parameters are $\theta = \begin{bmatrix} 5.3 & 1.1 & 8.45 & 2.35 \end{bmatrix}^T$. The history stack updates and the purging calculations are performed asynchronously in a separate thread, and the simulations are performed in real time.

The developed technique is compared against numerical differentiation-based concurrent learning where the numerical derivatives are computed using polynomial regression over a window of collected data. The state measurements are filtered using a moving average filter for state-derivative estimation. To facilitate the comparison, multiple simulation runs are performed using a combination of gains, window sizes, and thresholds for two levels of noise. The low-noise and high-noise simulations are performed by adding white Gaussian noise with variance 0.005 and 0.1, respectively, to the state measurements. The simulations are repeated five times for each combination of gains, and the gains, window sizes, and thresholds that yield the lowest steady-state RMS error over five runs are selected for comparison. Table I indicates that the developed technique outperforms numerical differentiation-based concurrent learning in both of the cases. The following figures illustrate the performance of the developed technique in one sample run with additive Gaussian noise with variance 0.005, a smoothing window of 80 samples, and the learning gains $k = 5$, $\gamma = 1$, $\mathcal{T} = 0.0019$, $\xi = 0.95$, $k_x = \text{diag}(11, 11, 91, 91)$, and $\alpha = 1$.

Figure 1a shows the evolution of the system state, where the added noise signal can be observed. Figure 1b demonstrates convergence of the unknown parameters to a neighborhood of their true values, where the dashed lines represent the true values. Figure 2a shows the convergence of the state estimation error to a ball around the origin. Figure 2b shows the convergence of the state-derivative estimation error to a ball around the origin. The transients in Figures 2a and 2b necessitate the need for history stack purging. Figure 3a

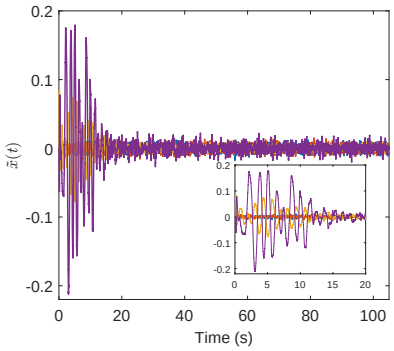


(a) State trajectory.

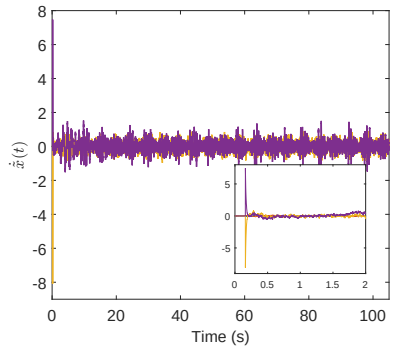


(b) Trajectories of the parameter estimates. Dashed lines represent the ideal values.

Fig. 1: Trajectories of the system state and the parameter estimates.



(a) State estimation error.



(b) State-derivative estimation error.

Fig. 2: Performance of the state-derivative estimator.

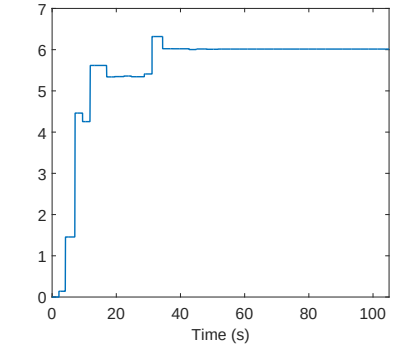
shows the minimum singular value of the history stack $\mathcal{H}$. The singular value is increasing because of the thresholding algorithm. In this simulation, the threshold parameter, ξ, is set to one. Figure 3b shows the increments of the purging index, It can be observed that the history stack gets purged faster initially as transients offer significant data, and then the rate of purging levels off approximately to a constant as the system achieves steady state.[9]

To demonstrate the sensitivity of the developed technique to changes in learning gains, a one-at-a-time sensitivity analysis is performed. To reduce the computational load, the gains of the high-gain derivative estimator are left unchanged. The parameters $\mathcal{T}$ and ξ, introduced in Algorithm 1, and the learning gain k, introduced in (2) are selected for the sensitivity analysis. The values of the steady-state RMS error for 7 different values of each parameter are presented in Table II. The results in Table II indicate that the developed method is robust to small changes in the learning gains.

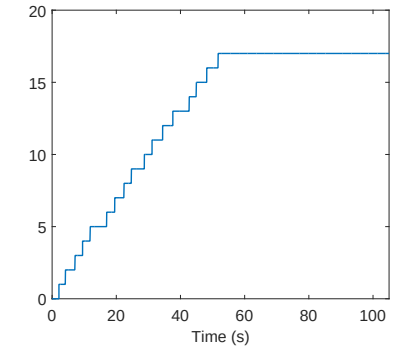
VI. Concluding Remarks

A concurrent learning-based parameter estimator is developed for a nonlinear system. An adaptive observer that employs full state feedback is employed to generate the

[9]The measurement noise does not inject excitation into the system since the noisy measurements are used only for parameter estimation, and the true state is used for feedback control.



(a) Minimum singular value of the history stack.



(b) Purging index

Fig. 3: Purging index and the minimum singular value for the history stack.

TABLE II: Sensitivity analysis results for the developed technique. The nominal values of k, $\mathcal{T}$, and ξ are selected to be $\xi = 0.95$, $\mathcal{T} = 0.0018$, and $k = 5$. The noise variance is set to 0.005 and the values of the remaining parameters are selected as $k_x = \text{diag}\,(11, 11, 91, 91)$, $\alpha = 1$, and $\gamma = 1$, and a moving average filter window of 80 samples is used to filter the state measurements.

$k =$	2	3	4	5	6	7	8
RMS steady-state error	0.61	0.39	0.30	0.31	0.33	0.46	0.57
$\xi =$	0.8	0.85	0.9	0.95	1	1.05	1.1
RMS steady-state error	0.39	0.36	0.29	0.31	0.32	0.36	0.49
$\mathcal{T} =$	0.0012	0.0014	0.0016	0.0018	0.0020	0.0022	0.0024
RMS steady-state error	0.35	0.38	0.36	0.31	0.33	0.32	0.36

state-derivative estimates required for concurrent learning. The developed technique is validated via simulations on a nonlinear system where the state measurements are corrupted by additive Gaussian noise. The simulation results indicate that the developed technique yields better results than numerical differentiation-based CL, even more so as the variance of the additive noise is increased. Even though the simulation results indicate a degree of robustness to measurement noise, the theoretical development does not account for measurement noise.

Measurement noise affects the developed parameter estimator in two ways. An error is introduced in the state-derivative estimates generated using the adaptive observer, and an error is introduced via the history stack since the state measurements recorded in the history stack are corrupted by noise. The former can be addressed if a noise rejecting observer such as a Kalman filter is used to generate the state-derivative estimates. The latter can be addressed by the use of an inherently noise-robust function approximation technique, e.g., a Gaussian process, to approximate the system dynamics. An extension of the developed parameter estimator that uses output feedback and is provably robust to measurement noise is a topic for future research.

REFERENCES

[1] N. Sureshbabu and J. Farrell, "Wavelet-based system identification for nonlinear control," *IEEE Trans. Autom. Control*, vol. 44, no. 2, pp. 412–417, 1999.

[2] O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media, 2001.

[3] H. Jaeger, "Tutorial on training recurrent neural networks, covering BPTT, RTRL, EKF and the "echo state network" approach," German National Research Center for Information Technology, Bremen, Tech. Rep. 159, 2002.

[4] P. P. Angelov and D. P. Filev, "An approach to online identification of takagi-sugeno fuzzy models," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 34, no. 1, pp. 484–498, 2004.

[5] M. A. Duarte and K. Narendra, "Combined direct and indirect approach to adaptive control," *IEEE Trans. Autom. Control*, vol. 34, no. 10, pp. 1071–1075, Oct 1989.

[6] M. Krstić, P. V. Kokotović, and I. Kanellakopoulos, "Transient-performance improvement with a new class of adaptive controllers," *Syst. Control Lett.*, vol. 21, no. 6, pp. 451–461, 1993.

[7] E. Lavretsky, "Combined/composite model reference adaptive control," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2692–2697, Nov 2009.

[8] G. V. Chowdhary and E. N. Johnson, "Theory and flight-test validation of a concurrent-learning adaptive controller," *J. Guid. Control Dynam.*, vol. 34, no. 2, pp. 592–607, Mar. 2011.

[9] H. Fukushima, T.-H. Kim, and T. Sugie, "Adaptive model predictive control for a class of constrained linear systems based on the comparison model," *Automatica*, vol. 43, no. 2, pp. 301–308, 2007.

[10] V. Adetola, D. DeHaan, and M. Guay, "Adaptive model predictive control for constrained nonlinear systems," *Syst. Control Lett.*, vol. 58, no. 5, pp. 320–326, 2009.

[11] G. Chowdhary, M. Mühlegg, J. How, and F. Holzapfel, "Concurrent learning adaptive model predictive control," in *Advances in Aerospace Guidance, Navigation and Control*, Q. Chu, B. Mulder, D. Choukroun, E.-J. van Kampen, C. de Visser, and G. Looye, Eds. Springer Berlin Heidelberg, 2013, pp. 29–47.

[12] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.

[13] P. Abbeel, M. Quigley, and A. Y. Ng, "Using inaccurate models in reinforcement learning," in *Int. Conf. Mach. Learn.* New York, NY, USA: ACM, 2006, pp. 1–8.

[14] D. Mitrovic, S. Klanke, and S. Vijayakumar, "Adaptive optimal feedback control with learned internal dynamics models," in *From Motor Learning to Interaction Learning in Robots*, ser. Studies in Computational Intelligence, O. Sigaud and J. Peters, Eds. Springer Berlin Heidelberg, 2010, vol. 264, pp. 65–84.

[15] M. P. Deisenroth and C. E. Rasmussen, "Pilco: A model-based and data-efficient approach to policy search," in *Int. Conf. Mach. Learn.*, 2011, pp. 465–472.

[16] R. Kamalapurkar, P. Walters, and W. E. Dixon, "Model-based reinforcement learning for approximate optimal regulation," *Automatica*, vol. 64, pp. 94–104, 2016.

[17] G. Chowdhary, "Concurrent learning for convergence in adaptive control without persistency of excitation," Ph.D. dissertation, Georgia Institute of Technology, Dec. 2010.

[18] G. Chowdhary, T. Yucelen, M. Mühlegg, and E. N. Johnson, "Concurrent learning adaptive control of linear systems with exponentially convergent bounds," *Int. J. Adapt. Control Signal Process.*, vol. 27, no. 4, pp. 280–301, 2013.

[19] M. Mühlegg, G. Chowdhary, and E. Johnson, "Concurrent learning adaptive control of linear systems with noisy measurements," in *Proc. AIAA Guid. Navig. Control Conf.*, 2012.

[20] A. Gelb, *Applied optimal estimation*. The MIT press, 1974.

[21] G. D. L. Torre, G. Chowdhary, and E. N. Johnson, "Concurrent learning adaptive control for linear switched systems," in *Proc. Am. Control Conf.*, Jun. 2013, pp. 854–859.

[22] B. Reish and G. Chowdhary, "Concurrent learning adaptive control for systems with unknown sign of control effectiveness," in *Proc. IEEE Conf. Decis. Control*, 2014, pp. 4131–4136.

[23] F. L. Lewis, S. Jagannathan, and A. Yesildirak, *Neural network control of robot manipulators and nonlinear systems*. Philadelphia, PA: CRC Press, 1998.

[24] A. Levant, "Robust exact differentiation via sliding mode technique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.

[25] ——, "Higher-order sliding modes, differentiation and output-feedback control," *Int. J. Control*, vol. 76, no. 9/10, pp. 924–941, 2003.

[26] L. Fridman, Y. Shtessel, C. Edwards, and X. G. Yan, "Higher-order sliding-mode observer for state estimation and input reconstruction in nonlinear systems," *Int. J. Robust Nonlinear Control*, vol. 18, pp. 399–412, 2007.

[27] L. Fridman, A. Levant, and J. Davila, *Observation and Identification Via High-Order Sliding Modes*. Springer Berlin / Heidelberg, 2008.

[28] P. A. Ioannou and P. V. Kokotovic, Eds., *Adaptive Systems with Reduced Models*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin Heidelberg, 1983, vol. 47, ch. 5. Adaptive control in the presence of disturbances, pp. 81–90.

[29] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.