

# A Novel Online Pseudospectral Method for Approximation of Nonlinear Systems Dynamics

Arian Yousefian<sup>1</sup>, Avimanyu Sahoo<sup>1</sup>, and Vignesh Narayanan<sup>2</sup>

**Abstract**—This note presents an online pseudospectral method for system identification using Chebyshev polynomial basis under aperiodic sampling. The system dynamics are approximated piecewise by introducing a sliding time window. The number of sampling instants (Chebyshev nodes) within each sliding window is selected dynamically based on a proposed node-selection criterion that guarantees desired approximation accuracy. The system states are measured at these aperiodic instants and used to estimate the coefficients of the basis polynomials using least squares. An adaptive state estimator is also proposed to reconstruct the continuous states using the approximated dynamics. The boundedness of the parameter and state estimation errors is proven analytically and validated numerically.

**Index Terms**—Chebyshev polynomial, least squares estimation, nonlinear system identification, online pseudospectral method.

## I. INTRODUCTION

System identification (ID) is fundamental to engineering systems because of unavailability of precise models due to complex and uncertain dynamics. It facilitates data-driven modeling essential for analysis, control, and decision making in applications such as robotics, aerospace vehicles, and power networks [1]. While linear ID is well studied, nonlinear ID remains challenging, especially online, where models are updated in real time [2]. Recent efforts on system ID span classical parametric schemes, spectral methods [3] and pseudospectral (PS) methods [4]. Spectral techniques rely on eigenfunction decompositions [5], whereas PS methods employ orthogonal polynomial collocation [6]. Classical orthogonal polynomial families (Legendre, Chebyshev, Hermite, Laguerre) underpin these methods (see, for example, [7]).

The PS methods, particularly those employing Chebyshev polynomials, are attractive due to minimax uniform approximation, node clustering that mitigates Runge oscillations, and efficient coefficient computation [7]–[9]. Most PS ID schemes, however, assume offline data over a fixed domain [10]. In parallel, neural networks (NNs) [11] and event-based NN training approaches [12] for online approximation have been developed. These approaches typically require higher sampling during learning [12] and the approximation accuracy depends on the number of neurons [11].

In our previous work [13], an online Chebyshev-based PS scheme with a sliding window was proposed where sampling

time instants were chosen at Chebyshev nodes, and state measurements at those instants were used to construct the basis. Because the state measurements were not the roots of the Chebyshev polynomial (Chebyshev nodes) of the basis, uniform approximation-error bounds were not guaranteed. To address this gap, this note introduces an online Chebyshev PS approach for ID that guarantees a uniform approximation with a desired accuracy under reduced, aperiodic sampling.

Instead of using states, the Chebyshev basis polynomial is constructed as a function of time over arbitrary intervals. The number of Chebyshev nodes (which determines the basis polynomial order and sampling points) within each sliding time window is selected adaptively to meet the desired accuracy. The coefficients of the basis function are estimated by least squares regression from the measured states and their derivatives at Chebyshev time nodes within each window, yielding a piecewise approximation of the system dynamics. The identified dynamics are then used to design an adaptive state estimator. To ensure continuity across sliding-windows, the approximation coefficients are recomputed so that the estimated dynamics match at each window transition point.

Furthermore, a detailed errors analysis and adaptive selection of the number of Chebyshev nodes, establishing a criterion for approximation order and sampling are presented. The boundedness of both the parameter and state estimation errors is proven analytically and, further, substantiated through simulation using a benchmark nonlinear oscillator system. This work builds on the preliminary work presented in [14]. The main contributions of this letter, beyond the preliminary work in [14], are: i) an online Chebyshev PS ID method; ii) an adaptive node-selection strategy with a theoretical guarantee on the desired approximation accuracy; iii) an adaptive PS state estimator; and iv) the convergence analysis establishing boundedness parameter and state estimation errors.

This note is organized as follows. Section II reviews offline PS methods and formulates the problem. Section III introduces the online PS framework, followed by Section IV, which discusses adaptive state estimation and node selection. Section V presents the simulation results to validate the analytical claims, followed by Conclusions in Section VI.

**Notations.**  $\mathbb{R}^n$  is the  $n$ -dimensional Euclidean space;  $\mathbb{R}^{n \times m}$  the set of  $n \times m$  real matrices.  $\|\cdot\|$  denotes the Euclidean norm. The superscript  $(\cdot)^T$  denotes transpose. For a matrix  $A$ ,  $\text{Sym}(A)$  denotes  $A + A^T$ . The ceiling and floor operators and Kronecker product are denoted by  $\lceil \cdot \rceil$ ,  $\lfloor \cdot \rfloor$ , and  $\otimes$ , respectively. If  $x(t) : \mathbb{R} \rightarrow \mathbb{R}^n$  is essentially bounded, then it is in the set  $\mathcal{L}_\infty^n := \{x(t) \mid \|x\|_\infty \triangleq \sup_{t \geq 0} \|x(t)\| < \infty\}$ , the set of all real-valued functions defined on the interval  $[a, b]$  that are  $N$ -

This work is partially supported by NSF #2327409, #2337998, and #2337999, and NSWC-PC #N00174-23-1-0006.

<sup>1</sup>Arian Yousefian and Avimanyu Sahoo are with the Electrical and Computer Engineering Department at the University of Alabama in Huntsville, Huntsville, AL (Emails: ay0024@uah.edu, and avimanyu.sahoo@uah.edu), and <sup>2</sup>Vignesh Narayanan is with the AI Institute, University of South Carolina (Email: vignar@sc.edu).

times continuously differentiable is denoted by  $C^N([a, b])$ .

## II. BACKGROUND ON CHEBYSHEV POLYNOMIAL-BASED APPROXIMATION AND PROBLEM STATEMENT

This section briefly reviews offline PS approximation with Chebyshev polynomial bases and formulates the problem.

### A. Chebyshev Approximation on the Standard Interval

Let  $g(x) \in C^{N+1}[-1, 1]$  be a scalar function. Using a first-kind Chebyshev PS basis  $T_i(x)$ ,  $g(x)$  admits the expansion

$$g(x) = \sum_{i=0}^{\infty} c_i^* T_i(x) = \sum_{i=0}^N c_i^* T_i(x) + \epsilon(x), \quad (1)$$

where  $\{c_i^*\}_{i=0}^N$  are the unknown coefficients and  $\epsilon(x)$  is the truncation error and  $\{T_i(x)\}_{i=0}^N$  on  $[-1, 1]$  are defined recursively as

$$\begin{cases} T_0(x) = 1, & T_1(x) = x, \\ T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x), & i = 2, 3, \dots, N. \end{cases} \quad (2)$$

The roots of  $T_i(x)$ , called the *Chebyshev nodes*, are widely used in polynomial interpolation. For  $T_N(x)$ , the nodes are

$$x_k = \cos((k - 0.5)\pi/N), \quad k = 1, 2, \dots, N. \quad (3)$$

The nodes are non-uniformly spaced and denser near the interval endpoints [9].

In practice, the domain often differs from  $[-1, 1]$ . Therefore, an affine map can be used to transform  $x$  in  $[-1, 1]$  to  $[a, b]$  as

$$x^S = \frac{1}{2}((a + b) + (b - a)x), \quad a \leq x^S \leq b. \quad (4)$$

Thus, the Chebyshev basis on  $[a, b]$  can be obtained by applying the change of variable in (4), which map it back to  $[-1, 1]$ , as

$$T_i^S(x^S) = T_i((2x^S - (a + b))/(b - a)). \quad (5)$$

Similarly, the nodes on  $[a, b]$  can be computed as

$$x_k^S = \frac{1}{2}((a + b) + (b - a)x_k), \quad k = 1, 2, \dots, N. \quad (6)$$

For brevity, hereafter, we use  $x$  for the shifted variable  $x^S$  and use the superscript  $S$  to indicate shifted polynomials.

### B. Chebyshev Expansion on a Shifted Interval

The function  $g(x)$  on  $[a, b]$  admits the shifted expansion  $g(x) = \sum_{i=0}^N c_i^{S*} T_i^S(x) + \epsilon(x)$ , where  $\{c_i^{S*}\}_{i=0}^N$  are the unknown coefficients in the shifted Chebyshev basis, and the PS approximation is given by

$$\hat{g}(x) = \sum_{i=0}^N c_i^S T_i^S(x), \quad (7)$$

where  $c_i^S$  denotes the estimates of  $c_i^{S*}$ .

To estimate  $c_i^S$ , one can sample  $y_k = g(x_k)$  at  $m \geq N + 1$  shifted Chebyshev nodes  $x_k \in [a, b]$  and employ least squares, given by [8]

$$c^S = (X^T X)^{-1} X^T Y, \quad (8)$$

where  $c^S = [c_0^S \ c_1^S \ \dots \ c_N^S]^T$ , and

$$X = \begin{pmatrix} T_0^S(x_1) & T_1^S(x_1) & \dots & T_N^S(x_1) \\ T_0^S(x_2) & T_1^S(x_2) & \dots & T_N^S(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ T_0^S(x_m) & T_1^S(x_m) & \dots & T_N^S(x_m) \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

From an interpolation error perspective, for each  $x \in [a, b]$  there exists a point  $\xi_x \in [a, b]$  such that [15]

$$g(x) - \hat{g}(x) = \frac{g^{(N+1)}(\xi_x)}{(N+1)!} \prod_{k=1}^{N+1} (x - x_k). \quad (9)$$

Since Chebyshev nodes minimize the maximum interpolation error, with these nodes, the error is bounded by

$$\max_{x \in [a, b]} |g(x) - \hat{g}(x)| \leq \frac{2D^{(N+1)}}{(N+1)!} \left(\frac{b-a}{4}\right)^{N+1}, \quad (10)$$

where  $D^{(N+1)} = \max_{\xi_x \in [a, b]} |g^{(N+1)}(\xi_x)|$ .

The PS methods are typically leveraged for offline learning using pre-collected data with known  $[a, b]$  and order  $N$ . However, online PS approximation introduces several challenges and are detailed below.

### C. Problem Formulation

Consider a continuous-time nonlinear system described by

$$\dot{x}(t) = F(x(t)), \quad (11)$$

where  $x(t) \in \mathbb{R}^{N_p}$  is the state vector and  $F : \mathcal{D} \rightarrow \mathbb{R}^{N_p}$  denotes the unknown nonlinear dynamics. The domain  $\mathcal{D} \subset \mathbb{R}^{N_p}$  contains the origin, and  $F(0) = 0$ . The function  $F$  is locally Lipschitz, and  $x(t) \in \mathcal{L}_{\infty}^{N_p}$ . The system in (11) can be considered as a closed-loop system or an autonomous system and the identification scheme can be used for prediction or control design.

The objective is to approximate  $F(x(t))$  online via the PS method, using Chebyshev polynomials of the first-kind as basis functions, with aperiodically sampled data at Chebyshev nodes to ensure the desired accuracy. Equivalently, we seek an approximation  $\hat{F}(x(t)) \in \mathcal{Q}_M$ , the set of degree- $M$  polynomials, that minimizes the interpolation error

$$\mathcal{E}(x(t)) = \min_{\hat{F}(x(t)) \in \mathcal{Q}_M} \left\| F(x(t)) - \hat{F}(x(t)) \right\|_{\infty}. \quad (12)$$

For  $x(t) \in \mathbb{R}^{N_p}$ , the dynamics (11) can be approximated as

$$\hat{F}(x(t)) = \eta^T \bigotimes_{j=1}^{N_p} \mathbb{T}_M^S(x^j(t)), \quad j = 1, 2, \dots, N_p, \quad (13)$$

for  $x^j(t) \in [a^j, b^j]$ , where  $\mathbb{T}_M^S(x^j(t)) = [T_0^S(x^j(t)), T_1^S(x^j(t)), \dots, T_M^S(x^j(t))]^T$  is the shifted Chebyshev basis vector,  $\eta \in \mathbb{R}^{(M+1)^{N_p} \times N_p}$  is the estimated-coefficient matrix [13], and  $M$  is the node count.

Computing Chebyshev nodes using (6) requires prior knowledge of the interval  $[a^j, b^j]$  and the node count  $M^j$  for each state  $x^j(t)$ . Since the dynamics are unknown, the bounds  $a^j, b^j$  and the required  $M^j$  cannot be specified a priori, and the nodes cannot be precomputed. Further, to measure the states

at the nodes, i.e.,  $\{x_k^j\}_{k=1}^{M^j+1}$ , requires the corresponding times  $\{t_k^j\}_{k=1}^{M^j+1}$ . Since there is no one-to-one map from state values to time, the sampling times cannot be determined for sampling and state measurement at nodes.

Therefore, for online Chebyshev PS approximation, the problem is fourfold: (i) to develop an online method that does not require sampling at state-space Chebyshev nodes; (ii) to determine suitable approximation intervals during operation to enable data collection; (iii) to select the polynomial degree or node count to meet a desired approximation accuracy; and (iv) to reconstruct the continuous state from aperiodic measurements. The next sections present a detailed solution.

### III. PROPOSED ONLINE PS FRAMEWORK

In this section, an online approximation framework based on the PS method is proposed using a sliding time window and the minimum number of nodes to guarantee desired approximation accuracy, as illustrated in Fig. 1.

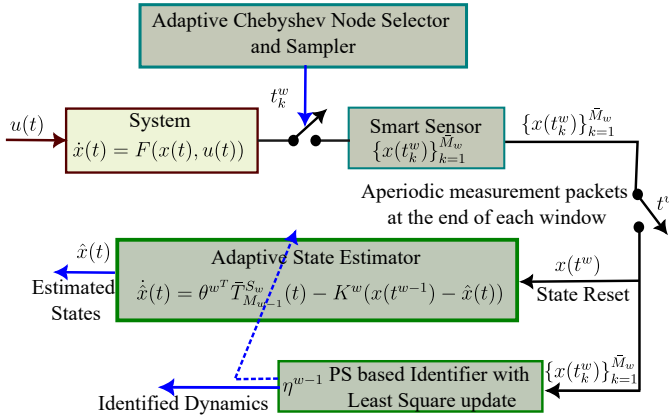


Fig. 1. Schematic of proposed approach.

Within each window, a smart sensor samples the system states at time nodes determined by the adaptive selector, packages the data, and transmits it to the identifier at the end of each window. The identifier performs batch least squares estimation to obtain coefficients (parameters) of the basis function. These identified parameters are then used to construct the adaptive state estimator that can estimate the system states from aperiodic sampling for the next window. The number and location of nodes, which determine both the sampling instants and the polynomial approximation order, are updated adaptively to ensure the desired approximation accuracy while minimizing computational effort.

#### A. Online PS Method for System Identification

The dynamics in (11) are an implicit function of time. Each component  $F^j(x(t))$ ,  $j = 1, \dots, N_p$ , can be approximated by the first-kind Chebyshev polynomial basis with time as the argument [10], given by

$$F^j(t) = \sum_{i=0}^M \eta_{ij}^* T_i(t) + \epsilon_j(t), \quad (14)$$

where  $\eta_{ij}^* \in \mathbb{R}$  are the unknown coefficients,  $\epsilon_j(t) \in \mathbb{R}$  is the truncation error, and  $M$  is approximation order (node count). The approximation of the function can be expressed as

$$\hat{F}^j(t) = \sum_{i=0}^M \eta_{ij} T_i(t), \quad (15)$$

where  $\eta_{ij} \in \mathbb{R}$  are the estimates of the target coefficients  $\eta_{ij}^*$ .

*Remark 1:* Constructing the Chebyshev basis as a function of time avoids using a high-dimensional state basis for  $N_p$ -dimensional functions in (13). This also allows online approximation and determining sampling points at time nodes.

To enable online approximation of the dynamics  $F(t)$ , define a sliding window  $(t^{w-1}, t^w]$   $w = 1, 2, \dots$ , where  $t^w$  denotes the transition between consecutive windows and  $\tau^w = t^w - t^{w-1}$  is the window width. Using (6), the  $M_w + 1$  time nodes (sampling instants) within  $(t^{w-1}, t^w]$  can be computed as

$$t_k^w = \frac{1}{2}(t^{w-1} + t^w) + \frac{1}{2}\tau^w \cos\left(\frac{k-0.5}{M_w+1}\pi\right), \quad (16)$$

for  $k = 1, 2, \dots, M_w + 1$ .

*Remark 2:* For ease of exposition, in this paper, we considered constant window width, i.e.,  $\tau^w = \tau$ , where  $\tau$  is a constant. A variable  $\tau^w$  can also be designed with the use of event-triggered or self-triggered control strategies [16], where the sampling intervals are adaptively determined based on stability and performance criteria in a controlled system.

With time nodes as in (16), the sensor samples  $x(t_k^w)$  at each Chebyshev node  $t_k^w$  in window  $w$  and also  $x(t_k^w - \Delta t)$  with  $\Delta t > 0$  to estimate  $\dot{x}(t)$  at the nodes. In addition, the state is measured at the start of each window to reset the identifier (discussed later in Sec. IV). All samples from the  $w$ -th window are used to estimate the coefficients at window transition instants.

*Remark 3:* For simplicity, state derivatives are approximated by a backward difference, which is known to amplify noise. To address this challenge, one can use adaptive or sliding mode state derivative estimators [7], [17].

From (14), the piecewise representation of the  $j$ -th system dynamics for  $j = 1, 2, \dots, N_p$  within  $w$ -th window is

$$F^j(t) = \sum_{i=0}^{M_w} \eta_{ij}^{w*} T_i^{S_w}(t) + \epsilon_j^w(t), \quad t^{w-1} < t \leq t^w, \quad (17)$$

where  $T_i^{S_w}(t)$  is the shifted Chebyshev polynomial on the window  $t^{w-1} < t \leq t^w$ . Defining the basis vector as

$$\bar{T}_{M_w}^{S_w}(t) = [T_0^{S_w}(t) \quad T_1^{S_w}(t) \quad \dots \quad T_{M_w}^{S_w}(t)]^\top, \quad (18)$$

the augmented system dynamics  $F^w(t) \in \mathbb{R}^{N_p}$  in (17) can be expressed as

$$F^w(t) = \eta^{w* \top} \bar{T}_{M_w}^{S_w}(t) + \epsilon^w, \quad t^{w-1} < t \leq t^w, \quad (19)$$

where  $\eta^{w*} = [\eta_1^{w*} \quad \eta_2^{w*} \quad \dots \quad \eta_{N_p}^{w*}] \in \mathbb{R}^{(M_w+1) \times N_p}$  with  $\eta_j^{w*} \in \mathbb{R}^{(M_w+1)}$  is the parameter matrix for the truncated Chebyshev polynomial. The approximation of (19) on  $w$ -th window can be expressed as

$$\hat{F}^w(t) = \eta^{w \top} \bar{T}_{M_w}^{S_w}(t), \quad t^{w-1} < t \leq t^w, \quad (20)$$

where  $\eta^w \in \mathbb{R}^{(M_w+1) \times N_p}$  is the estimated coefficients.

Using the  $w$ -th window measurements at  $\{t_k^w\}_{k=1}^{M_w+1}$ , the least-squares estimator in (8) provides  $\eta^w$  at  $t^w$  as

$$\eta^w = (\bar{T}_{M_w}^{S_w}(t) \bar{T}_{M_w}^{S_w \top}(t) + R_0^w)^{-1} (R_0^w \eta_0^w + \bar{T}_{M_w}^{S_w}(t) \dot{X}^w(t)), \quad (21)$$

where  $\bar{T}_{M_w}^{S_w}(t) \in \mathbb{R}^{(M_w+1) \times (M_w+1)}$  and  $\dot{X}^w \in \mathbb{R}^{(M_w+1) \times N_p}$  with

$\bar{T}_{M_w}^{S_w}(t) = [\bar{T}_{M_w}^{S_w}(t_1^w) \quad \bar{T}_{M_w}^{S_w}(t_2^w) \cdots \bar{T}_{M_w}^{S_w}(t_{(M_w+1)}^w)]$  and  $\dot{X}^w = [\dot{x}(t_1^w) \quad \dot{x}(t_2^w) \cdots \dot{x}(t_{(M_w+1)}^w)]^\top$  for  $t^{w-1} < t \leq t^w$ . Matrices  $R_0^w$  and  $\eta_0^w$  are user-defined to ensure the first-term inverse in (21) exists [18].

### B. Minimum Nodes for Desired Approximation Accuracy

The key advantage of piecewise approximation with a moving window is the ability to dynamically choose the degree of the polynomial  $M_w$  to keep the approximation accuracy within a desired bound in each window. The following theorem provides the criteria for selecting the minimum node count using the maximum error per-window. Before introducing the node selection criteria, the following lemma is necessary.

*Lemma 1:* Define  $\mathcal{E}_{\max}^w = \max_{t^{w-1} < t \leq t^w} \|F^w(t) - \hat{F}^w(t)\|$ . For the dynamics (19) and its piecewise Chebyshev approximation (20) on  $(t^{w-1}, t^w]$ , it holds that

$$\mathcal{E}_{\max}^w \leq C_w \alpha_{m,w}^{-M_w}, \quad (22)$$

where  $0 < C_w < \frac{2D_w^{(M_w+1)}}{\sqrt{6\pi}} \left(\frac{\tau}{4}\right)^{M_w+1}$ ,  $M_w \geq 2$ , and  $\alpha_{m,w} = \frac{M_w+1}{e}$  with Euler's number  $e \approx 2.718$ .

*Proof:* The proof is given in the Appendix.

*Theorem 1:* Let  $F^w(t)$  in (19) be piecewise continuous on  $(t^{w-1}, t^w]$ , where the window length satisfies  $\tau < 4$ . Given a desired approximation error  $\epsilon_{\text{th}}$  and an initial node count  $M_w$ , and using the approximation  $\hat{F}^w(t)$  in (20), the minimum node count that keeps the  $\mathcal{E}_{\max}^w$  within  $[\kappa\epsilon_{\text{th}}, \epsilon_{\text{th}}]$  can be computed as

$$\bar{M}_w = \begin{cases} M_w + \lceil \rho_1^w \ln(\mathcal{E}_{\max}^w / \epsilon_{\text{th}}) \rceil, & \mathcal{E}_{\max}^w > \epsilon_{\text{th}}, \\ M_w, & \kappa\epsilon_{\text{th}} \leq \mathcal{E}_{\max}^w \leq \epsilon_{\text{th}}, \\ M_w + \lceil \rho_2^w \ln(\mathcal{E}_{\max}^w / \kappa\epsilon_{\text{th}}) \rceil, & \mathcal{E}_{\max}^w < \kappa\epsilon_{\text{th}}, \end{cases} \quad (23)$$

where  $0 < \kappa < 1$ ,  $\rho_1^w = \frac{1}{\ln(\alpha_{m,w})}$ ,  $\alpha_{m,w} = \frac{M_w+1}{e}$ ,  $\rho_2^w = \frac{1}{\ln(\frac{1}{\alpha_{m,w}})}$ ,  $\alpha_{m,w} = \frac{M_w+1}{e}$ , and  $\bar{M}_w$  is the node count satisfying  $\mathcal{E}_{\max}^w > \kappa\epsilon_{\text{th}}$ .

*Proof:* The proof is given in the Appendix.

*Remark 4:* Given a tolerance  $\epsilon_{\text{th}} \in \mathbb{R}_{>0}$ , an error band  $[\kappa\epsilon_{\text{th}}, \epsilon_{\text{th}}]$  with  $0 < \kappa < 1$  is employed to create a dead zone. When the error lies within this band, the node count is held constant.

*Remark 5:* The node count in (23) uses a conservative error bound  $C_w \alpha_{m,w}^{-M_w}$  (see (A.10) in the Appendix), which is below the threshold  $\epsilon_{\text{th}}$ . Therefore, (23) is a sufficient condition only and the actual node count needed to meet  $\epsilon_{\text{th}}$  may be less than or equal to the node count  $\bar{M}_w$ .

## IV. ADAPTIVE STATE ESTIMATOR AND NODE DESIGN

This section presents the design of an adaptive state estimator that leverages approximated system dynamics and a node selection criterion to ensure accurate state estimation.

### A. Adaptive State Estimator Design

Unlike traditional ID, the proposed online PS approach approximates the dynamics piecewise as in (20). Therefore, the state estimator dynamics can only be defined piecewise for each window  $w$ . However, the coefficients  $\eta^w$  for  $w$ -th window are computed at the end of the window at time  $t^w$  using (21) from the data collected between  $(t^{w-1}, t^w]$ . Hence, the estimated coefficients are not available for implementation in the  $w$ -th window. Using the available information, i.e., previous window's coefficients  $\eta^{w-1}$  and number of Chebyshev nodes  $M_{w-1}$ , the state estimator dynamics for the  $w$ -th window can be defined as

$$\dot{\hat{x}}(t) = \eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_w}(t) - K^w (x(t^{w-1}) - \hat{x}(t)), \quad (24)$$

for  $t^{w-1} < t \leq t^w$ , where  $K^w \in \mathbb{R}^{N_p \times N_p}$  represents the gain matrix for window  $w$ . Note that  $\eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_w}(t)$  is the dynamics computed using available estimated parameters  $\eta^{w-1}$  from  $(w-1)$ -th window and shifted basis for  $w$ -th window (super script  $S_w$ ) of order  $M_{w-1}$ .

Although, the state estimator in (24) is implementable, guaranteeing stability for all  $t \in \mathbb{R}_{\geq 0}$  may be challenging due to the use of a piecewise approximated dynamics, which may lack sufficient smoothness at the window transitions. In addition, since the shifted Chebyshev polynomials from any window map their domains to  $[-1, 1]$ , the dynamics  $\eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_w}(t)$  in (24) is a delayed version of the approximated dynamics in the previous window ( $\eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_{w-1}}(t)$ ), as demonstrated in the following proposition, which can further introduce errors in state estimation.

*Proposition 1:* Consider the piecewise approximation of the system dynamics in (20) on  $(t^{w-2}, t^{w-1}]$ , given by  $\hat{F}^{w-1}(t) = \eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_{w-1}}(t)$ , and on  $(t^{w-1}, t^w]$  by  $\hat{F}^w(t) = \eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_w}(t)$ . For equal-width windows  $\tau$ , it holds that  $\hat{F}^{w-1}(t) = \hat{F}^w(t + \tau)$ .

*Proof:* The proof is given in the Appendix.

To address these challenges, the continuity of the approximated dynamics and their derivatives must be enforced at the transition between adjacent windows. This requires updating the coefficients  $\eta^{w-1}$  corresponding to the  $w$ -th window.

Let the new approximated dynamics for  $w$ -th window of the state estimator using  $M_{w-1}$  nodes are given by

$$\hat{F}_\theta^w(t) = \theta^{w \top} \bar{T}_{M_{w-1}}^{S_w}(t), \quad t^{w-1} < t \leq t^w, \quad (25)$$

where  $\theta^w \in \mathbb{R}^{M_{w-1}+1}$  are the updated coefficients. For continuity of the dynamics and its derivatives up to order  $M_{w-1}$  at  $t^{w-1}$ , it must hold that

$$\left. \frac{d^p}{dt^p} (\eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_{w-1}}(t)) \right|_{t=t^{w-1}} = \left. \frac{d^p}{dt^p} (\theta^{w \top} \bar{T}_{M_{w-1}}^{S_w}(t)) \right|_{t=t^{w-1}}, \quad (26)$$

for  $p = 0, \dots, M_{w-1}$ . Since the left side of (26) is known,  $\theta^w$  can be computed by solving the simultaneous equations.

*Lemma 2:* Given the continuity condition stated in (26), the solution  $\theta^w$  exists and is unique.

*Proof:* The proof is given in the Appendix.

With (25), the adaptive state estimator dynamics in (24) on  $(t^{w-1}, t^w]$  can be expressed as

$$\dot{\hat{x}}(t) = \theta^{w \top} \bar{T}_{M_{w-1}}^{S_w}(t) - K^w (x(t^{w-1}) - \hat{x}(t)), \quad (27)$$

where

$$\hat{x}(t) = \begin{cases} x(t^{w-1}) & t = t^{w-1}, \\ \hat{x}(t) & t^{w-1} < t \leq t^w. \end{cases} \quad (28)$$

At the start of each window, the state estimates in (27) are reset to the measured values via (28).

While the using coefficient vector  $\theta^w$  enables online state estimation, it inevitably introduces additional approximation error, as demonstrated in the following theorem.

*Theorem 2:* Consider the dynamics (11), approximated by the Chebyshev expansion (25). Let  $\theta^w$  be computed according to (26). Then the function approximation error  $\mathcal{E}_\theta^w(t) = F^w(t) - \hat{F}_\theta^w(t)$  for  $t \in (t^{w-1}, t^w]$  is bounded above as

$$|\mathcal{E}_\theta^w(t)| \leq \frac{2D_{w-1}^{(M_{w-1}+1)}}{M_{w-1}!} \left(\frac{\tau}{4}\right)^{M_{w-1}+1} + L^w \tau + 2\sqrt{M_{w-1}+1}\|\theta^w\|, \quad (29)$$

where  $L^w$  is the Lipschitz constant on the  $w$ -th time window.

*Proof:* The proof is given in the Appendix.

From Theorem 2, the approximation error splits into three parts. The first is the same as the offline Chebyshev bound (10). The second and third terms quantify the effect of window length and online parameter estimation errors. To minimize the  $\mathcal{E}_\theta^w(t)$  from available data, we define its windowed average and then redefine the node-selection criterion.

### B. Adaptive Node Selection Approach

The error  $\mathcal{E}_\theta^w(t)$  is not computable for all  $t$ , since  $F^w(t)$  is unknown. With the available information at nodes  $t_k^w$ , the average approximation error on  $(t^{w-1}, t^w]$  can be defined as

$$\mathcal{E}_{\theta, \text{avg}}^w = \frac{1}{M_w + 1} \sum_{k=1}^{M_w+1} \left\| F^w(t_k^w) - \hat{F}_\theta^w(t_k^w) \right\|. \quad (30)$$

*Remark 6:* The approximation error at node points is zero when  $\hat{\mathcal{F}}^w(t)$  is obtained from (20), i.e.,  $\mathcal{F}^w(t_k^w) = \hat{\mathcal{F}}^w(t_k^w)$ . Since the state estimator utilizes (25), the error in (30) is nonzero unless the approximation is exact.

The average approximation error  $\mathcal{E}_{\theta, \text{avg}}^w$  in (30) can be compared with a desired approximation error  $\epsilon_{\text{th}}$  at each window to dynamically adjust the node count for the next window. Recalling Theorem 1, the adaptive update law for the node count that keeps the  $\mathcal{E}_{\theta, \text{avg}}^w$  within  $[\kappa\epsilon_{\text{th}}, \epsilon_{\text{th}}]$  can be expressed as

$$M_{w+1} = \begin{cases} M_w + \left\lceil \gamma_1^w \ln \left( \mathcal{E}_{\theta, \text{avg}}^w / \epsilon_{\text{th}} \right) \right\rceil, & \mathcal{E}_{\theta, \text{avg}}^w > \epsilon_{\text{th}}, \\ M_w, & \kappa\epsilon_{\text{th}} \leq \mathcal{E}_{\theta, \text{avg}}^w \leq \epsilon_{\text{th}}, \\ M_w + \left\lceil \gamma_2^w \ln \left( \mathcal{E}_{\theta, \text{avg}}^w / \kappa\epsilon_{\text{th}} \right) \right\rceil, & \mathcal{E}_{\theta, \text{avg}}^w < \kappa\epsilon_{\text{th}}, \end{cases} \quad (31)$$

where, to account for the system nonlinearity, parameters  $0 < \gamma_1^w, \gamma_2^w < 10$  are set per the proof of Theorem 1.

*Remark 7:* According to Remark 5, (31) swaps the rounding operators in (23). This also prevents round-off errors from adding an unnecessary node. Constants  $\rho_1^w, \rho_2^w$  are replaced with adaptive gains  $\gamma_1^w, \gamma_2^w$ , which provide additional degrees of freedom to account for the rapid change in slope of the nonlinear dynamics.

The next lemma and theorem present the boundedness analysis for the parameter and state estimation errors.

*Lemma 3:* Consider the dynamics (11) on  $(t^{w-1}, t^w]$  and their approximation in (25). Let  $\theta^w$  is computed using (26). Then the parameter estimation error  $\tilde{\theta}^w = \theta^{w*} - \theta^w$  is ultimately bounded (UB).

*Proof:* The proof is given in the Appendix.

*Theorem 3:* Consider the dynamics (11) on  $(t^{w-1}, t^w]$ , approximated by the Chebyshev expansion (25), and the adaptive state estimator (27). Let  $\theta^w$  be updated according to (26). Then the state-estimation error  $\tilde{x}$  and the parameter-estimation error  $\tilde{\theta}^w$  are UB, provided that the gain matrix  $K^w$  satisfies the Lyapunov equation

$$Z^w K^w + K^{w\top} Z^w = -Q^w, \quad (32)$$

where  $Z^w, Q^w$  are symmetric positive definite matrices, provided that  $\lambda_{\min}(Q^w) > 3$  and  $\|\tilde{x}\|^2 > \Upsilon^w / (\lambda_{\min}(Q^w) - 3)$ , where  $\Upsilon^w$  is a constant defined in (A.39).

*Proof:* The proof is given in the Appendix.

## V. SIMULATION RESULTS

This section presents numerical results using the the Stuart–Landau oscillator, given by

$$\begin{aligned} \dot{x}_1 &= (a - (x_1^2 + x_2^2))x_1 - \omega x_2, \\ \dot{x}_2 &= (a - (x_1^2 + x_2^2))x_2 + \omega x_1, \end{aligned} \quad (33)$$

with  $a = 0.5$  and  $\omega = 1.5$ . The system exhibits a stable limit cycle of radius  $\sqrt{a}$ . The simulation parameters were selected as follows: windows length  $\tau = 0.2$  s, desired approximation error  $\epsilon_{\text{th}} = 10^{-3}$ , time horizon 12 s, and sampling period 0.001 s. The initial approximation order was chosen as  $M_w = 2$  and updated using (30) and (31) with  $\gamma_1^w = 0.2$ ,  $\gamma_2^w = 0.9$ , and  $\kappa = 0.1$ . The system and identifier states were initialized as  $[0.5 \ 0.5]^\top$  and  $[2 \ 2]^\top$ , respectively, and the identifier was reset in each window using (28). The gains  $K^w$  followed (32) with  $Z^w = 10I_2$  and  $Q^w = \text{diag}(5, 4.5)$ . Initial coefficients for the first window were chosen as  $\eta^1 = [0.05 \ \mathbf{1}_{3 \times 1} \ -0.05 \ \mathbf{1}_{3 \times 1}]$  where  $\mathbf{1}_{3 \times 1} = [1 \ 1 \ 1]^\top$ .

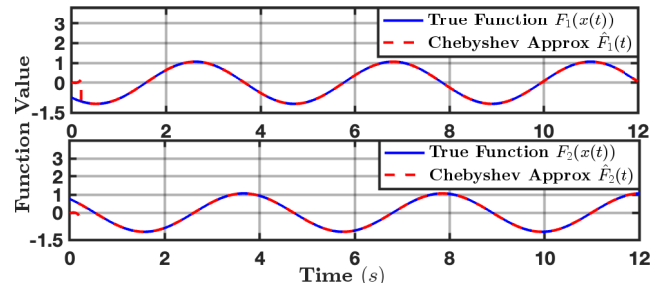


Fig. 2. Convergence of the Chebyshev approximation  $\hat{F}(t)$  to the true dynamics  $F(x(t))$ .

Fig. 2 shows that the approximated dynamics converge to the true dynamics by 0.4 s (window 2). The first window shows a large error due to the initial condition and an insufficient node count (two), which is inadequate to approximate the dynamics. As the node count gets updated in later windows, the error decreases.

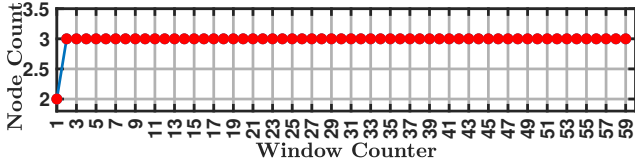


Fig. 3. Number of nodes in each time window for approximation.

Fig. 3 shows the adaptive adjustment of node count across moving time windows. Starting from arbitrary initial conditions, the algorithm increases the count to match local complexity, reaching 3 by the second window and driving the error to  $10^{-3}$ . As the function behavior tends to become nearly linear within each window, the error falls below  $10^{-3}$ . In concave regions it increases again near  $10^{-3}$ . With  $\kappa = 0.1$ , since the error remains within  $[\kappa\epsilon_{th}, \epsilon_{th}]$ , the node count stays constant, demonstrating the method's effectiveness in guaranteeing the desired approximation accuracy. Therefore, compared to neural network-based approximations that rely on periodic sampling (total 12,000 samples), the proposed aperiodic approach offers a significant reduction in sampling frequency (only 35 samples) and, therefore, computational effort.

Figures 4 and 5 depict, respectively, the true state with its estimate  $\hat{x}(t)$  and the estimation error  $\tilde{x}(t)$ . As shown, the estimates converge to the true states within three windows. The adaptive estimator meets the desired accuracy under aperiodic sampling, which validates the gain design in (32) and the parameter convergence.

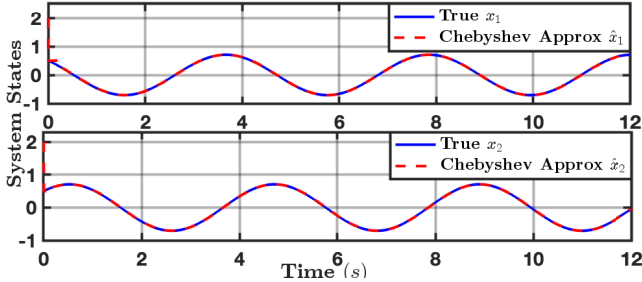


Fig. 4. Convergence of Chebyshev approximation  $\hat{x}(t)$  to system state  $x$ .

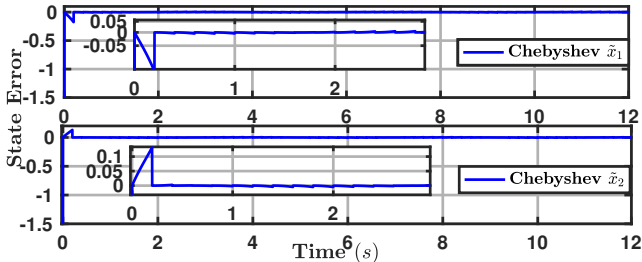


Fig. 5. Convergence of the state estimation error  $\tilde{x}(t)$ .

## VI. CONCLUSIONS

This note, presented an online PS method for system ID. The results demonstrate that the proposed online ID framework using first-kind Chebyshev polynomials can accurately approximate nonlinear system dynamics with only a few aperiodic

samples per window. The adaptive estimator effectively reconstructed continuous states, and the extended Lyapunov analysis confirmed convergence and boundedness of the parameter and state estimation errors. Numerical studies further validated the analytical findings. This approach can be employed to design controller by jointly approximating the drift and input vector fields and is included in our future work.

## APPENDIX

**Proof of Lemma 1:** In (10), the factorial term obeys the extended Stirling double inequality [19], so it holds that

$$\sqrt{2\pi M_w} \left(\frac{M_w}{e}\right)^{M_w} e^{\frac{1}{12M_w+1}} < M_w! < \sqrt{2\pi M_w} \left(\frac{M_w}{e}\right)^{M_w} e^{\frac{1}{12M_w}}. \quad (\text{A.1})$$

Taking reciprocals in (A.1) for  $M_w + 1$  yields

$$\frac{1}{(M_w + 1)!} < \frac{1}{\sqrt{2\pi(M_w + 1)}} \left(\frac{e}{M_w + 1}\right)^{M_w + 1} e^{-\frac{1}{12(M_w + 1) + 1}}. \quad (\text{A.2})$$

From (A.2), for  $M_w \geq 2$ , the following inequalities hold

$$\begin{aligned} \frac{1}{\sqrt{2\pi(M_w + 1)}} &\leq \frac{1}{\sqrt{6\pi}}, \quad e^{-\frac{1}{12(M_w + 1) + 1}} < 1, \\ \left(\frac{e}{M_w + 1}\right)^{M_w + 1} &< \alpha^{-(M_w + 1)}, \quad \text{for } \alpha < \frac{M_w + 1}{e}, \\ \alpha^{-(M_w + 1)} &< \alpha^{-(M_w)}, \quad \text{for } \alpha > 1. \end{aligned} \quad (\text{A.3})$$

From (10), with  $\alpha_{m,w} = \sup \alpha = \frac{M_w + 1}{e}$  and using (A.3), the maximum approximation-error is bounded by

$$\mathcal{E}_{\max}^w \leq C_w \alpha_{m,w}^{-M_w}, \quad t^{w-1} < t \leq t^w, \quad (\text{A.4})$$

where  $M_w \geq 2$ , and  $C_w$  satisfying

$$0 < C_w < \frac{2D_w^{(M_w + 1)}}{\sqrt{6\pi}} \left(\frac{\tau}{4}\right)^{(M_w + 1)}. \quad (\text{A.5})$$

This completes the proof.  $\blacksquare$

**Proof of Theorem 1:** Based on (23), the proof is carried out in three cases.

**Case I:** For the case  $\mathcal{E}_{\max}^w > \epsilon_{th}$ . To reduce the error below  $\epsilon_{th}$ , an increment  $\Delta M_w = \bar{M}_w - M_w > 0$  in  $w$ -th window is required such that

$$\bar{C}_w \alpha_{m,w}^{-(M_w + \Delta M)} < \epsilon_{th}, \quad t^{w-1} < t \leq t^w, \quad (\text{A.6})$$

where  $\bar{M}_w$  is the minimum number of nodes, and  $\bar{C}_w$  satisfies

$$0 < \bar{C}_w < \frac{2D_w^{(M_w + \Delta M_w + 1)}}{\sqrt{6\pi}} \left(\frac{\tau}{4}\right)^{(M_w + \Delta M_w + 1)}. \quad (\text{A.7})$$

Defining  $\bar{\alpha}_{m,w} = (\bar{M}_w + 1)/e$  and since  $\bar{M}_w > M_w$  it holds that

$$\bar{\alpha}_{m,w} > \alpha_{m,w}. \quad (\text{A.8})$$

From (A.6) and (A.8) it can be verified that

$$\bar{C}_w \bar{\alpha}_{m,w}^{-\bar{M}_w} < \bar{C}_w \alpha_{m,w}^{-(M_w + \Delta M)} < \epsilon_{th}. \quad (\text{A.9})$$

Using (A.9) and (A.4), and denoting  $\bar{\mathcal{E}}_{\max}^w$  as the updated maximum error after node increment, leads to

$$\bar{\mathcal{E}}_{\max}^w \leq \bar{C}_w \bar{\alpha}_{m,w}^{-\bar{M}_w} < \epsilon_{th}. \quad (\text{A.10})$$

Since  $\mathcal{E}_{\max}^w > \epsilon_{\text{th}}$ , by (A.4) and (A.10) one can get

$$\bar{C}_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)} < C_w \alpha_{m,w}^{-M_w}. \quad (\text{A.11})$$

Since  $C_w \alpha_{m,w}^{-M_w} > C_w \alpha_{m,w}^{-(M_w + \Delta M_w)}$ , two sub-cases arise.

**1) Sub-case I.1:** For the case

$$\bar{C}_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)} \leq C_w \alpha_{m,w}^{-(M_w + \Delta M_w)}. \quad (\text{A.12})$$

Given (A.5) and (A.7), the condition  $\frac{\tau}{4} < 1$  makes  $(\frac{\tau}{4})^{M_w+1}$  and  $(\frac{\tau}{4})^{\bar{M}_w+1}$  very small, making  $C_w$  and  $\bar{C}_w$  tend to zero. So, from (A.9) and (A.12) one can write

$$\bar{C}_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)} \leq C_w \alpha_{m,w}^{-(M_w + \Delta M_w)} < \epsilon_{\text{th}}. \quad (\text{A.13})$$

Using (A.13), (A.10) satisfies  $\bar{\mathcal{E}}_{\max}^w \leq C_w \alpha_{m,w}^{-\bar{M}_w} < \epsilon_{\text{th}}$ . Then taking logarithms both sides yield

$$\Delta M_w \geq \frac{1}{\ln(\alpha_{m,w})} [\ln(C_w/\epsilon_{\text{th}}) - M_w \ln(\alpha_{m,w})]. \quad (\text{A.14})$$

From (A.4), dividing by  $\epsilon_{\text{th}}$  and taking logarithm, it holds

$$\ln(\mathcal{E}_{\max}^w/\epsilon_{\text{th}}) \leq \ln(C_w/\epsilon_{\text{th}}) - M_w \ln(\alpha_{m,w}). \quad (\text{A.15})$$

Thus, rearranging (A.15) and substituting to (A.14) leads to

$$\Delta M_w \geq \frac{1}{\ln(\alpha_{m,w})} \ln(\mathcal{E}_{\max}^w/\epsilon_{\text{th}}). \quad (\text{A.16})$$

Since  $M_w \geq 2$ ,  $\alpha_{m,w} \leq \frac{3}{e}$  and  $0 < \frac{1}{\ln(\alpha_{m,w})} \leq 10.15$ . So, denoting  $\rho_1^w = \frac{1}{\ln(\alpha_{m,w})}$ , the number of integer nodes can be given as

$$\bar{M}_w = M_w + \lceil \rho_1^w \ln(\mathcal{E}_{\max}^w/\epsilon_{\text{th}}) \rceil. \quad (\text{A.17})$$

**2) Sub-case I.2:** For this sub-case, it holds that  $\bar{C}_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)} > C_w \alpha_{m,w}^{-(M_w + \Delta M_w)}$ . So, using (A.8), the following inequality holds

$$\bar{C}_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)} > C_w \alpha_{m,w}^{-(M_w + \Delta M_w)} > C_w \bar{\alpha}_{m,w}^{-(M_w + \Delta M_w)}. \quad (\text{A.18})$$

From (A.18),  $\bar{C}_w > C_w$ . Hence, (A.4) satisfies

$$\mathcal{E}_{\max}^w \leq \bar{C}_w \alpha_{m,w}^{-M_w}, \quad t^{w-1} < t \leq t^w. \quad (\text{A.19})$$

From (A.9) and (A.10),  $\bar{\mathcal{E}}_{\max}^w \leq \bar{C}_w \alpha_{m,w}^{-\bar{M}_w} < \epsilon_{\text{th}}$ . Taking the logarithm of the right-hand side yields

$$\Delta M_w \geq \frac{1}{\ln(\alpha_{m,w})} [\ln(\bar{C}_w/\epsilon_{\text{th}}) - M_w \ln(\alpha_{m,w})]. \quad (\text{A.20})$$

From (A.19), dividing by  $\epsilon_{\text{th}}$  and taking logarithm, we have

$$\ln(\mathcal{E}_{\max}^w/\epsilon_{\text{th}}) \leq \ln(\bar{C}_w/\epsilon_{\text{th}}) - M_w \ln(\alpha_{m,w}). \quad (\text{A.21})$$

From (A.20) and (A.21), we can obtain  $\Delta M_w \geq \frac{1}{\ln(\alpha_{m,w})} \ln(\mathcal{E}_{\max}^w/\epsilon_{\text{th}})$ . Therefore, the minimum number of nodes is same as in (A.17).

**Case II:** For this case  $\kappa \epsilon_{\text{th}} \leq \mathcal{E}_{\max}^w \leq \epsilon_{\text{th}}$ . This case is trivial since the maximum error is within the desired error bound and the number of nodes  $M_w$  is the minimum.

**Case III:** For this case  $\mathcal{E}_{\max}^w < \kappa \epsilon_{\text{th}}$ . In this case, nodes are reduced until  $\kappa \epsilon_{\text{th}}$ . Denote by  $\underline{M}_w$  the node count on the last interval with  $\mathcal{E}_{\max}^w \geq \kappa \epsilon_{\text{th}}$ . When  $\mathcal{E}_{\max}^w \leq \kappa \epsilon_{\text{th}}$ , applying (A.16), the change in node count satisfies

$$\Delta M_w \geq \rho_2^w \ln(\kappa \epsilon_{\text{th}}/\mathcal{E}_{\max}^w), \quad (\text{A.22})$$

where  $\rho_2^w = \frac{1}{\ln(\underline{\alpha}_{m,w})}$  and  $\underline{\alpha} = \frac{M_w+1}{e}$ . In (A.22),  $\Delta M_w$  is the minimum node-count decrement to bring  $\mathcal{E}_{\max}^w$  to  $\kappa \epsilon_{\text{th}}$ . So, the minimum number of nodes  $\bar{M}_w = M_w - \Delta M_w$  is

$$\bar{M}_w = M_w + \lfloor \rho_2^w \ln(\mathcal{E}_{\max}^w/\kappa \epsilon_{\text{th}}) \rfloor. \quad (\text{A.23})$$

This completes the proof.  $\blacksquare$

**Proof of Proposition 1:** Define  $t$  on  $(t^{w-2}, t^{w-1}]$  and  $\bar{t}$  on  $(t^{w-1}, t^w]$ . With equal interval length  $\tau$ , we have  $t = \bar{t} - \tau$ .

From (20), the approximation on  $(t^{w-1}, t^w]$  using  $\eta^{w-1}$  (estimated on  $(t^{w-2}, t^{w-1})$ ) is

$$\hat{F}^w(\bar{t}) = \eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_w}(\bar{t}), \quad t^{w-1} < \bar{t} \leq t^w. \quad (\text{A.24})$$

Using (18) and (5), on  $(t^{w-2}, t^{w-1}]$  the basis for window  $w-1$  can be expressed as

$$\bar{T}_{M_{w-1}}^{S_{w-1}}(t) = \bar{T}_{M_{w-1}} \left( \frac{2t - (t^{w-2} + t^{w-1})}{t^{w-1} - t^{w-2}} \right). \quad (\text{A.25})$$

Similarly, for the interval  $(t^{w-1}, t^w]$ , i.e., window  $w$ , the basis

$$\bar{T}_{M_{w-1}}^{S_w}(\bar{t}) = \bar{T}_{M_{w-1}} \left( \frac{2\bar{t} - (t^{w-1} + t^w)}{t^w - t^{w-1}} \right). \quad (\text{A.26})$$

Using  $\bar{t} = t + \tau$ ,  $\tau = t^{w-1} - t^{w-2}$ ,  $t^w = t^{w-2} + 2\tau$ , (A.26) leads to

$$\bar{T}_{M_{w-1}}^{S_w}(\bar{t}) = \bar{T}_{M_{w-1}} \left( \frac{2t - (t^{w-2} + t^{w-1})}{\tau} \right). \quad (\text{A.27})$$

From (A.27) and (A.25),  $\bar{T}_{M_{w-1}}^{S_w}(\bar{t}) = \bar{T}_{M_{w-1}}^{S_{w-1}}(t)$ . Left-multiplying by  $(\eta^{w-1})^\top$  and invoking (A.24) gives

$$\hat{F}^w(\bar{t}) = \hat{F}^w(t + \tau) = \eta^{w-1 \top} \bar{T}_{M_{w-1}}^{S_{w-1}}(t) = \hat{F}^{w-1}(t). \quad (\text{A.28})$$

This completes the proof.  $\blacksquare$

**Proof of Lemma 2:** Stacking (26) for  $p=0, \dots, M_{w-1}$  gives

$$\begin{aligned} & \left[ \begin{array}{cccc} T_0^{S_w}(t) & T_1^{S_w}(t) & \cdots & T_{M_{w-1}}^{S_w}(t) \\ 0 & \frac{dT_1^{S_w}(t)}{dt} & \cdots & \frac{dT_{M_{w-1}}^{S_w}(t)}{dt} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{d^{M_{w-1}} T_{M_{w-1}}^{S_w}(t)}{dt^{M_{w-1}}} \end{array} \right] \Bigg|_{t=t^{w-1}} \left[ \begin{array}{c} \theta_{0j}^w \\ \theta_{1j}^w \\ \vdots \\ \theta_{M_{w-1}j}^w \end{array} \right] \\ & = \left[ \begin{array}{c} \eta_j^{w-1 \top} \bar{T}_{M_{w-1}}^{S_{w-1}}(t) \\ \eta_j^{w-1 \top} \frac{d\bar{T}_{M_{w-1}}^{S_{w-1}}(t)}{dt} \\ \vdots \\ \eta_j^{w-1 \top} \frac{d^{M_{w-1}} \bar{T}_{M_{w-1}}^{S_{w-1}}(t)}{dt^{M_{w-1}}} \end{array} \right] \Bigg|_{t=t^{w-1}}, \quad j = 1, 2, \dots, N_p. \end{aligned} \quad (\text{A.29})$$

With (5) and  $T_i^{S_w}(t^{w-1}) = T_i(-1)$ , each derivative in (A.29) can be expressed as

$$\frac{d^p T_i^{S_w}(t)}{dt^p} \Bigg|_{t=t^{w-1}} = \left( \frac{2}{\tau} \right)^p \frac{d^p T_i(t)}{dt^p} \Bigg|_{t=-1}. \quad (\text{A.30})$$

By [20, Eq. (3.35)], the  $p$ -th derivative of  $T_i(t)$  at  $t = -1$  admits a non-zero closed-form expression. So from (A.30), it holds  $\frac{d^p T_i^{S_w}(t)}{dt^p} \Big|_{t=t^{w-1}} \neq 0$ . Hence the upper triangular matrix in (A.29) has a nonzero diagonal and is invertible, so a unique  $\theta^w$  exists. This completes the proof.  $\blacksquare$

**Proof of Theorem 2:** Recalling (25), the function approximation error can be written as

$$\begin{aligned} \mathcal{E}_\theta^w(t) &= F^w(t) - \hat{F}_\theta^w(t) = (F^w(t) - F^w(t^{w-1})) + (F^w(t^{w-1}) \\ &\quad - \hat{F}_\theta^w(t^{w-1})) + (\hat{F}_\theta^w(t^{w-1}) - \hat{F}_\theta^w(t)), t^{w-1} < t \leq t^w. \end{aligned} \quad (\text{A.31})$$

With  $|t - t^{w-1}| < \tau$  on  $(t^{w-1}, t^w]$ , the first term in (A.31) satisfies

$$\|F^w(t) - F^w(t^{w-1})\| \leq L^w |t - t^{w-1}| \leq L^w \tau. \quad (\text{A.32})$$

From (26) for  $p = 0$ , we have  $\hat{F}^{w-1}(t^{w-1}) = \hat{F}_\theta^w(t^{w-1})$ , and trivially  $F^{w-1}(t^{w-1}) = F^w(t^{w-1})$ . Thus, using (10), the second term in (A.31) is given by

$$\|(F^w(t^{w-1}) - \hat{F}_\theta^w(t^{w-1}))\| \leq \frac{2D_{w-1}^{(M_{w-1}+1)}}{M_{w-1}!} \left(\frac{\tau}{4}\right)^{M_{w-1}+1}, \quad (\text{A.33})$$

and  $D_{w-1}^{(M_{w-1}+1)} = \max_{\xi \in [t^{w-2}, t^{w-1}]} \|F^{(M_{w-1}+1)}(\xi)\|$  is a bounded scalar.

Finally, using (25), the third term in (A.31) is given by

$$\hat{F}_\theta^w(t^{w-1}) - \hat{F}_\theta^w(t) = \theta^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t^{w-1}) - \theta^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t). \quad (\text{A.34})$$

Recall that  $\|\bar{T}_{M_{w-1}}^{S_w}(t^{w-1})\| = \sqrt{M_{w-1}+1}$  and  $\|\bar{T}_{M_{w-1}}^{S_w}(t)\| \leq \sqrt{M_{w-1}+1}$ . Using the inequality  $\|A - B\| \leq \|A\| + \|B\|$ , (A.34) can be written as

$$\begin{aligned} \|\hat{F}_\theta^{w-1}(t^{w-1}) - \hat{F}_\theta^{w-1}(t)\| &\leq \|\theta^w\| \left( \|\bar{T}_{M_{w-1}}^{S_w}(t^{w-1})\| \right. \\ &\quad \left. + \|\bar{T}_{M_{w-1}}^{S_w}(t)\| \right) \leq 2\sqrt{M_{w-1}+1} \|\theta^w\|. \end{aligned} \quad (\text{A.35})$$

By combining (A.32), (A.33), and (A.35), the error bound in (29) can be reached. This completes the proof.  $\blacksquare$

**Proof of Lemma 3:** By the Chebyshev property, all entries of  $\bar{T}_{M_w}^{S_w}(t)$  lie in  $[-1, 1]$  for  $t \in (t^{w-1}, t^w]$ . Hence, (26) gives

$$R_0^w \leq \bar{T}_{M_w}^{S_w}(t) \bar{T}_{M_w}^{S_w\top}(t) + R_0^w < (M_w + 1)^2 I_{(M_w+1)} + R_0^w. \quad (\text{A.36})$$

It can be verified from (A.36) that  $\left(\bar{T}_{M_w}^{S_w}(t) \bar{T}_{M_w}^{S_w\top}(t) + R_0^w\right)^{-1} \leq (R_0^w)^{-1}$ . Thus, the inverse in (21) is bounded. From stability of (11),  $\dot{X}^w(t)$  is bounded. Hence  $\eta^{w-1}$  from (21) and  $\theta^w$  from (26) are bounded. Since the actual coefficients  $\theta^{w*}$  is bounded. Therefore  $\hat{\theta}^w = \theta^{w*} - \theta^w$  is also bounded. This completes the proof.  $\blacksquare$

**Proof of Theorem 3:** The state measurement error due to aperiodic availability of the state at the identifier can be defined as  $e_s^w = x^{w-1} - x$ . Since the system states are bounded,  $e_s^w$  is also bounded, i.e.,  $\|e_s^w\| < e_{s, M_w}$ . It holds that

$$\begin{aligned} \dot{\hat{x}}(t) &= \tilde{\theta}^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t) + K^w (x(t^{w-1}) - \hat{x}) + \bar{\epsilon}^w \\ &= \tilde{\theta}^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t) + K^w (e_s^w + \tilde{x}) + \bar{\epsilon}^w, t^{w-1} < t \leq t^w. \end{aligned} \quad (\text{A.37})$$

Consider the Lyapunov function candidate over  $t^{w-1} < t \leq t^w$  as  $\mathcal{V}^w = \tilde{x}^\top Z^w \tilde{x} + \frac{1}{2} \tilde{\theta}^{w\top} \tilde{\theta}^w$ .

Recall that  $\theta^w$  and  $\hat{\theta}^w$  are constant on  $(t^{w-1}, t^w]$ , hence  $\tilde{\theta}^w$  is constant and  $\dot{\tilde{\theta}}^w = 0$ . Thus, the Lyapunov first derivative is

$$\begin{aligned} \dot{\mathcal{V}}^w &= \text{Sym}(\tilde{x}^\top Z^w \tilde{\theta}^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t)) + \text{Sym}(\tilde{x}^\top Z^w K^w e_s^w) \\ &\quad + \text{Sym}(\tilde{x}^\top Z^w \bar{\epsilon}^w) + \tilde{x}^\top (Z^w K^w + K^{w\top} Z^w) \tilde{x}. \end{aligned} \quad (\text{A.38})$$

By Young's inequality,  $\text{Sym}(X^\top Y) \leq X^\top X + Y^\top Y$ ; hence

$$\text{Sym}(\tilde{x}^\top Z^w \tilde{\theta}^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t)) \leq \|\tilde{x}\|^2 + \|Z^w \tilde{\theta}^{w\top} \bar{T}_{M_{w-1}}^{S_w}(t)\|^2,$$

$\text{Sym}(\tilde{x}^\top Z^w K^w e_s^w) \leq \|\tilde{x}\|^2 + \|Z^w K^w e_s^w\|^2$ , and  $\text{Sym}(\tilde{x}^\top Z^w \bar{\epsilon}^w) \leq \|\tilde{x}\|^2 + \|Z^w \bar{\epsilon}^w\|^2$ . Recall  $\|\bar{\epsilon}^w\| \leq \epsilon_M^w$ ,  $\|\bar{T}_{M_{w-1}}^{S_w}(t)\| \leq 1$  and by Lemma 3,  $\|\hat{\theta}^w\| \leq \epsilon_\eta^w$ . So with (32), the Lyapunov first derivative is bounded by

$$\dot{\mathcal{V}}^w \leq -(\lambda_{\min}(Q^w) - 3) \|\tilde{x}\|^2 + \Upsilon^w, \quad (\text{A.39})$$

where  $\Upsilon^w = \|Z^w \epsilon_\eta^w\|^2 + \|Z^w \epsilon_M^w\|^2 + \|Z^w K^w e_s^w\|^2$ . From (A.39),  $\dot{\mathcal{V}}^w < 0$  as long as  $\lambda_{\min}(Q^w) > 3$  and  $\|\tilde{x}\|^2 > \frac{\Upsilon^w}{\lambda_{\min}(Q^w) - 3}$ . Therefore,  $\tilde{x}$  is UB with the bound  $\|\tilde{x}\|^2 \leq \frac{\Upsilon^w}{\lambda_{\min}(Q^w) - 3}$ . Further, since the estimated state is reset to the true state at each window transition, we have  $\tilde{x}(t^w) = 0$  at all window transition times, and it is bounded. Therefore, it follows that  $\hat{x}(t)$  is UB. This completes the proof.  $\blacksquare$

## REFERENCES

- [1] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*, 2nd ed. Springer, 2020.
- [2] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, 2010.
- [3] Y. Jin, L. Hou, and S. Zhong, "Extended dynamic mode decomposition with invertible dictionary learning," *Neural networks*, vol. 173, p. 106177, 2024.
- [4] A. Pirastehzad and M. J. Yazdanpanah, "A successive pseudospectral-based approximation of the solution of regulator equations," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1760–1775, 2021.
- [5] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, pp. 309–325, 2005.
- [6] P. G. Constantine, M. S. Eldred, and E. T. Phipps, "Sparse pseudospectral approximation method," *Computer Methods in Applied Mechanics and Engineering*, vol. 229, pp. 1–12, 2012.
- [7] L. N. Trefethen, *Approximation theory and approximation practice, extended edition*. SIAM, 2019.
- [8] J. C. Mason and D. C. Handscomb, *Chebyshev polynomials*. CRC press, 2002.
- [9] T. J. Rivlin, *Chebyshev polynomials*. Courier Dover Publications, 2020.
- [10] H. Ma, T. Qin, and W. Zhang, "An efficient Chebyshev algorithm for the solution of optimal control problems," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 675–680, 2010.
- [11] F. Lewis, S. Jagannathan, and A. Yesildirak, *Neural network control of robot manipulators and non-linear systems*. CRC press, 2020.
- [12] A. Sahoo, H. Xu, and S. Jagannathan, "Adaptive neural network-based event-triggered control of single-input single-output nonlinear discrete-time systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 1, pp. 151–164, 2015.
- [13] A. Yousefian, A. Sahoo, and V. Narayanan, "Aperiodic sensing and data-driven system identification of nonlinear systems using Chebyshev pseudospectral approach," in *Proc. of the IEEE 63rd Conference on Decision and Control (CDC)*, Millan, Italy, 2024, pp. 8766–8771.
- [14] —, "A novel online system identification using Chebyshev polynomial basis of the second kind," in *Proc. of the 2025 American Control Conference (ACC)*, Denver, CO, USA, 2025, pp. 1808–1813.
- [15] G. W. Stewart, *Afternotes on numerical analysis*. SIAM, 1996.
- [16] A. Sahoo, H. Xu, and S. Jagannathan, "Neural network-based event-triggered state feedback control of nonlinear continuous-time systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 3, pp. 497–509, 2015.
- [17] A. Levant, "Higher-order sliding modes, differentiation and output-feedback control," *International Journal of Control*, vol. 76, no. 9–10, pp. 924–941, 2003.
- [18] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and adaptive control design*. John Wiley & Sons, Inc., 1995.
- [19] W. Feller, *An introduction to probability theory and its applications, Volume 2*. John Wiley & Sons, 1991, vol. 2.
- [20] D. Johnson, "Chebyshev polynomials in the spectral tau method and applications to eigenvalue problems," Los Alamos National Laboratory, Tech. Rep., 1996.