Stabilization of nonlinear systems with unknown delays via delay-adaptive neural operator approximate predictors

Luke Bhan[a,*], Miroslav Krstic[a], Yuanyuan Shi[a]

[a] *University of California, San Diego, USA*

Abstract

This work establishes the first rigorous stability guarantees for approximate predictors in delay-adaptive control of nonlinear systems, addressing a key challenge in practical implementations where exact predictors are unavailable. We analyze two scenarios: (i) when the actuated input is directly measurable, and (ii) when it is estimated online. For the measurable input case, we prove semi-global practical asymptotic stability with an explicit bound proportional to the approximation error ϵ. For the unmeasured input case, we demonstrate local practical asymptotic stability, with the region of attraction explicitly dependent on both the initial delay estimate and the predictor approximation error. To bridge theory and practice, we show that neural operators—a flexible class of neural network-based approximators—can achieve arbitrarily small approximation errors, thus satisfying the conditions of our stability theorems. Experiments on two nonlinear benchmark systems—a biological protein activator/repressor model and a micro-organism growth Chemostat model—validate our theoretical results. In particular, our numerical simulations confirm stability under approximate predictors, highlight the strong generalization capabilities of neural operators, and demonstrate a substantial computational speedup of up to 15× compared to a baseline fixed-point method.

1 Introduction

In this work, we extend the concept of approximate predictors using operator learning to delay systems with constant but *unknown* delays. The presence of unknown delays presents a significant challenge, as the approximation of the predictor introduces an additive error requiring analysis akin to robust adaptive control. Furthermore, unlike robust adaptive control, the adaptive parameter estimated is the actuator delay which directly impacts system stability. Despite these challenges, we achieve results similar to [9], ensuring the semi-global practical stability of the feedback system, which depends on the approximation error and an additional error term on the delay bounds. This represents the first result for implementing any type of approximate predictor when the delay is constant but unknown.

* Corresponding author

Email addresses: `lbhan@ucsd.edu` (Luke Bhan), `krstic@ucsd.edu` (Miroslav Krstic), `yyshi@ucsd.edu` (Yuanyuan Shi).

[1] *Acknowledgements:* The work of L. Bhan is supported by the U.S. Department of Energy (DOE) grant DE-SC0024386. The work of M. Krstic was funded by AFOSR grant FA9550-23-1-0535 and NSF grant ECCS-2151525. The work of Y. Shi is supported by Department of Energy grant DE-SC0025495, Schmidt Sciences AI2050 Early Career Fellowship, and NSF grant ECCS-2442689.

1.1 Predictor feedback designs and implementations

Predictor feedback methods for compensating actuator delays in dynamical systems have been studied for over 50 years [50,19,29,21,22,42,23]. These approaches have proven effective across diverse applications such as traffic control [43,2], aerospace vehicles [55,17], and robotics [6,4,5]. However, in nonlinear systems, implementation remains challenging because the predictor is defined by an implicit ordinary differential equation [3]. To address this, [24] proposed numerical schemes combining finite differencing with successive approximations to approximate predictors for both linear and nonlinear systems, along with stability guarantees. Yet, this work does not consider delay-adaptive cases, leaving stability for predictors with unknown delays unestablished. Moreover, these schemes are computationally expensive due to the need for fine discretization. Recently, [9] introduced operator learning to approximate the predictor mapping, achieving speedups of 1000× compared to traditional solvers, but this was limited to nonlinear systems with constant, known delays. In this work, we extend [9] to handle unknown delays, addressing a critical gap in delay-adaptive predictor feedback.

While exact predictor feedback for unknown delays was initially addressed in [12], introducing an approximate predictor significantly complicates the analysis. Now, additive errors arise in the Lyapunov bounds, which still permit an ISS-like (input-to-state stability) result when the distributed input is known. However, when both the actuator input and delay are unknown, additional multiplicative approximation errors emerge, restricting stability to a smaller, local region of attraction depending both on the neural approximation error and the initial delay estimation error. This is similar to the results of robust adaptive control [46,45]- although significantly more challenging due to the adaptive parameter being the delay on the actuator. Nonetheless, we show that this region persists as long as the approximation errors remain sufficiently small **and** the initial delay estimation error is close enough.

1.2 Operator learning in control

As discussed above, because the predictor is implicitly defined, it must be *approximated*. In this work, we adopt a learning-based approach using neural operators for two reasons. First, neural operators, introduced in [13] and later popularized for learning PDE solution operators in [38,39], have become a powerful framework for approximating implicit operators in control. They have achieved speedups of up to 10^3 over traditional PDE solvers in PDE backstepping settings [10,31,54,53,48], and have also been applied in gain scheduling [34], adaptive control [33,11], and practical problems such as traffic flow [56] and oil drilling [52]. Second, operator-learning frameworks naturally support universal approximation results for ODE and PDE solution operators, making them well suited to the present theoretical study under Lipschitz regularity. For these reasons, neural operators offer a strong framework for predictor approximation, both analytically and computationally in real-time implementation.

1.3 Contributions

To summarize, the paper makes the following contributions:

(1) *Lipschitz regularity and universal approximation of delay-dependent predictors.* We prove Lipschitz continuity of the predictor operator and use it to establish uniform neural-operator approximation on compact sets.
(2) *Rigorous stability guarantees for delay-adaptive nonlinear systems under approximate predictor feedback.* For measurable actuator inputs, we prove semi-global practical asymptotic stability; for unmeasured actuator inputs, we prove a local practical stability result. In both cases, the bounds depend explicitly on the predictor approximation and delay-estimation errors.
(3) *Numerical validation on biologically motivated nonlinear systems.* We validate the theory on two benchmark systems and show that learned predictors preserve stability while providing substantial computational speedups over numerical predictor implementations.

1.4 Notation

For functions, $h : [0, D] \times \mathbb{R} \to \mathbb{R}$, we use $h_x(x,t) = \frac{\partial h}{\partial x}(x,t)$ and $h_t(x,t) = \frac{\partial h}{\partial t}(x,t)$ to denote derivatives.We use $C^k(\mathcal{S}_1, \mathcal{S}_2)$ to denote the set of functions with continuous k derivatives mapping sets $\mathcal{S}_1$ to $\mathcal{S}_2$. For a n-vector, we use $|\cdot|$ for the Euclidean norm. For functions, we define the spatial L^p norms as $\|h(\cdot,t)\|_{L^p(0,D)} = (\int_0^D |h(x,t)|^p dx)^{\frac{1}{p}}$ for $p \in [1, \infty)$ and may omit the domain to $\|h(\cdot,t)\|_{L^p}$ when clear. For brevity, we use the shorthand $\alpha_{i:j} \in \mathcal{K}_\infty$ to indicate that the functions $\alpha_i, \alpha_{i+1}, \ldots, \alpha_j$ are all elements of $\mathcal{K}_\infty$. For sets Ω, let $\mathrm{cl}(\Omega)$ denote its closure and $\overline{\Omega}$ is diameter. For a function U and parameter $D > 0$, define the historical control operator as $(T_D(t)U)(x) := U(t - D(x-1))$.

A preliminary version appeared at CDC 2025 [7]; this journal version adds theory for unmeasured actuator input (Section 6) and Chemostat simulations (Section 7).

2 Technical background

2.1 Delay-adaptive control

We study the nonlinear plant

$$\dot{X}(t) = f(X(t), U(t-D)), \tag{1}$$

where $X \in \mathbb{R}^n$, $U \in \mathbb{R}$, $f \in C^2(\mathbb{R}^n \times \mathbb{R}; \mathbb{R}^n)$ such that $f(0,0) = 0$, and D is an unknown delay within the interval $[\underline{D}, \overline{D}]$ where $\underline{D} > 0$. Note, the use of a single input U is made for simplicity of exposition. The approach and associated guarantees extend directly to the multi-input case with a common delay. As standard in predictor feedback designs [12], we require mild assumptions:

Assumption 1 *The plant $\dot{X} = f(X, \Omega)$ with Ω scalar is strongly forward complete.*

Assumption 2 *There exists $\kappa \in C^2(\mathbb{R}^n; \mathbb{R})$ such that the feedback law $U(t) = \kappa(X(t))$ guarantees that the delay-free plant is globally exponentially stable. Additionally, there exists constants $M_1, M_2 > 0$ such that the control law satisfies the growth conditions*

$$|\kappa(X)| \le M_1|X|, \qquad \left|\frac{d\kappa}{dX}(X)\right| \le M_2, \tag{2}$$

Notice by Assumption 2 and [26, Theorem 4.14] , we have that there exists $\lambda, C_1, C_2 > 0$ and a class C^∞ radially unbounded positive definite function V such that $\forall X \in \mathbb{R}^n$,

$$\frac{dV}{dX}(X) f(X, \kappa(X)) \le -\lambda V(X), \tag{3}$$

$$|X|^2 \le V(X) \le C_1 |X|^2, \tag{4}$$

$$\left|\frac{dV}{dX}(X)\right| \le C_2|X|. \tag{5}$$

Third, since $f \in C^2$, we have it is locally Lipschitz:

Assumption 3 *Let $f(X,U)$ be as in (1) and $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}$ be compact domains with bounds $\overline{X}$, $\overline{U}$ respectively. Then, there exists a constant $M_3(\overline{X}, \overline{U}) > 0$ such that f satisfies $\forall X_1, X_2 \in \mathcal{X}$ and $\forall u_1, u_2 \in \mathcal{U}$*

$$|f(X_1,u_1) - f(X_2,u_2)| \le M_3(|X_1 - X_2| + |u_1 - u_2|)\,, \tag{6}$$

Due to Assumption 3, there exists $M_4 > 0$ such that

$$\left|\frac{\partial f}{\partial X}(X,U)\right| \le M_4\,, \quad \forall (X,U) \in \mathcal{X} \times \mathcal{U}. \tag{7}$$

We emphasize that these assumptions are needed in the study of nonlinear systems with unknown, arbitrary length input delays. Assumption 1 guarantees that solutions of (1) exist for all forward time and do not escape before the delayed input $U(t-D)$ reaches the plant. Assumption 2 is used to obtain the semi-global result in Section 3 and may be relaxed to local exponential stability for a corresponding local result. Lastly, Assumption 3 imposes the mild regularity condition of local Lipschitz continuity. Consequently, the framework applies to nonlinear systems with smooth dynamics and known delay-free stabilizing controllers, including rigid-body dynamics such as robotic manipulators [6], adaptive cruise control systems [1], and power system flow dynamics [16].

As standard in predictor feedback designs, we reformulate the plant (1) into the following ODE-PDE cascade where the delay is absorbed into the transport PDE:

$$\dot{X}(t) = f(X(t), u(0,t))\,, \tag{8a}$$
$$Du_t(x,t) = u_x(x,t)\,, \qquad (x,t) \in (0,1) \times \mathbb{R}_+ \tag{8b}$$
$$u(1,t) = U(t)\,. \tag{8c}$$

The representation of the delay through a transport PDE was first introduced in [30] where the analytical solution of the transport PDE recovers the original plant. Namely, we have that

$$u(x,t) = U(t + D(x-1))\,, \quad x \in [0,1]\,, \tag{9}$$

and thus when $x = 0$ in the PDE representation, we recover $U(t-D)$ as in (1). To compensate for the delay, we estimate the state D timesteps in the future with the predictor given for all $x \in [0,1]$

$$p(x,t) = X(t+Dx) = X(t) + D\int_0^x f(p(y,t), u(y,t))dy\,, \tag{10}$$

which is itself, an implicit ODE whose analytical form is unknown for most nonlinear plant dynamics f. That is $p(1,t)$ is the *estimate* of the future state: $X(t+D)$. is Then, under the assumption that the delay is known, the stabilizing control law is given by $U(t) = \kappa(p(1,t))$. When the delay is unknown, as in the problem setting of this work, the control law is chosen by

$$U(t) = \kappa(\breve{p}(1,t))\,, \tag{11}$$

where the predictor is given by

$$\breve{p}(x,t) = X(t) + \breve{D}(t)\int_0^x f(\breve{p}(y,t), u(y,t))dy\,. \tag{12}$$

Following [12], $\breve{D}(t)$ is estimated as

$$\dot{\breve{D}}(t) = \gamma \mathrm{Proj}_{[\underline{D},\overline{D}]}\left\{\breve{D}(t), \phi(t)\right\}\,, \tag{13}$$

$$\phi(t) = -\frac{\int_0^1 (1+x)q_1(x,t)w(x,t)dx}{1 + V(x) + b\int_0^1 (1+x)w(x,t)^2 dx}\,, \tag{14}$$

where b is a user-specified parameter, V is the positive definite Lyapunov function in (3), (4), (5), w is given by the backstepping transformation

$$w(x,t) = u(x,t) - \kappa(\breve{p}(x,t))\,, \tag{15}$$

and the scalar function q_1 is defined as

$$q_1(x,t) = \frac{d\kappa}{d\breve{p}}(\breve{p}(x,t))\Phi(x,0,t)f(\breve{p}(0,t), u(0,t))\,, \tag{16}$$

where Φ is the transition matrix associated with the space-varying time-parameterized equation $(dr/dx)(x) = \breve{D}(t)(\partial f/\partial \breve{p})(\breve{p}(x,t), u(x,t))r(x)$. The projection is the standard projection operator given by

$$\mathrm{Proj}_{[\underline{D},\overline{D}]}\{\breve{D}, \phi\} = \begin{cases} 0, & \breve{D} = \overline{D},\ \phi > 0, \\ 0, & \breve{D} = \underline{D},\ \phi < 0, \\ \phi, & \text{otherwise.} \end{cases} \tag{17}$$

The update law $\dot{\breve{D}}(t)$ is not heuristic; it follows from the Lyapunov-based certainty-equivalent adaptive backstepping design for the transport-PDE representation (8a)–(8c) [12]. In particular, differentiating the backstepping transformation (15) introduces a q_1-dependent term involving the transition matrix Φ, and the law in (14) is chosen so that its numerator cancels this term in the Lyapunov derivative while its denominator provides the usual normalization.

Under the exact feedback law (11), [12, Theorem 1] gives global asymptotic stability. However, as mentioned, (12) is not analytically known and therefore needs to be approximated for any real-world implementation. Thus, we introduce an approximation for the predictor $\breve{p}$ and show that, under the universal approximation of neural operators, semi-global practical asymptotic stability is achieved.

2.2 Neural operators

For the reader's convenience, we briefly review the neural-operator background used later in the analysis. A neural operator is a neural-network-based approximation of a nonlinear operator acting between function spaces. In contrast to standard neural networks, which map one finite-dimensional vector to another, a neural operator is designed to learn mappings between input and output functions in a manner that is, in principle, discretization invariant (See [27] for a tutorial). For the analysis that follows, we will refer to a neural operator as any architecture satisfying the non-locality framework in [35, Section 1.2] which includes the most common variants (Fourier Neural Operator (FNO) [38] and DeepONet [39]). Under this framework, we have the following universal neural operator approximation theorem as follows:

Theorem 1 *[35, Theorem 1 and Corollary 1] Let $\Omega_u \subset \mathbb{R}^{d_{u_1}}$ and $\Omega_v \subset \mathbb{R}^{d_{v_1}}$ be two bounded domains with Lipschitz boundaries. Let $s, s' \geq 0$ be integers and $\Psi : C^s(\mathrm{cl}(\Omega_u)\,; \mathbb{R}^{d_{u_1}}) \to C^{s'}(\mathrm{cl}(\Omega_v)\,; \mathbb{R}^{d_{v_1}})$ be a continuous operator and fix a compact set $K \subset C^s(\mathrm{cl}(\Omega_u)\,; \mathbb{R}^{d_{u_1}})$. Then for any $\epsilon > 0$, there exists a neural operator $\hat{\Psi}$ (as in [35, Section 1.2]), such that for all $y \in \Omega_v$,*

$$\sup_{u \in K} |\Psi(u)(y) - \hat{\Psi}(u)(y)| \leq \epsilon\,. \tag{18}$$

We emphasize two key requirements of this theorem: the operator mapping is continuous and the existence of a neural operator only occurs on a arbitrary large *compact* function space. Hence, for universal approximation, we need to present Lipschitz regularity of the predictor operator. This is the focus of the following section.

3 Neural operator approximate predictors

Definition 1 *Let $X \in \mathbb{R}^n$, $U \in C^2([0,1]; \mathbb{R})$, $\varphi \in \mathbb{R}^+$. Then, we define the* ***predictor operator*** *as the mapping $\breve{\mathcal{P}} : (X, U, \varphi) \to \breve{P}$ where $\breve{P} = \breve{\mathcal{P}}(X, U, \varphi)$ satisfies for all $s \in [0,1]$,*

$$\breve{P}(s) - X - \varphi \int_0^s f(\breve{P}(\theta), U(\theta))d\theta = 0\,. \tag{19}$$

Notice that, by definition the predictor operator yields the solution to (12) where φ, as an input to the operator, is the estimated delay. To invokte Theorem 1, we prove:

Lemma 1 *Let Assumption 3 hold. Then, for any $X_1, X_2 \in \mathcal{X}$, $U_1, U_2 \in C^2([0,1]; \mathcal{U})$, and $\varphi_1, \varphi_2 \in (\underline{D}, \overline{D})$ the predictor operator $\breve{\mathcal{P}}$ given in Definition 1 satisfies*

$$\begin{aligned} &\|\breve{\mathcal{P}}(X_1, U_1, \varphi_1) - \breve{\mathcal{P}}(X_2, U_2, \varphi_2)\|_{L^\infty(0,1)} \\ &\qquad \leq C_{\breve{\mathcal{P}}} \left(|X_1 - X_2| + \|U_1 - U_2\|_{L^\infty(0,1)} + |\varphi_1 - \varphi_2|\right)\,, \end{aligned} \tag{20}$$

with Lipschitz constant

$$C_{\breve{\mathcal{P}}} = e^{\overline{D}M_3} \max\left\{1, \Xi, \overline{D}M_3\right\}\,, \tag{21}$$

$$\Xi = M_3 \left[\overline{U} + e^{\overline{D}M_3}(\overline{X} + M_3\overline{DU})\right]\,. \tag{22}$$

PROOF. For all $s \in [0,1]$, let $\bar{P}_1(s) := \breve{\mathcal{P}}(X_1, U_1, \varphi_1)$ and likewise $\bar{P}_2(s) := \breve{\mathcal{P}}(X_2, U_2, \varphi_2)$. Note that, for $s \in [0,1]$ the predictor is uniformly bounded. First, by definition and Assumption 3, we have

$$\begin{aligned} \bar{P}_1(s) &= X_1 + \varphi_1 \int_0^s f(\bar{P}_1(\theta), U_1(\theta))d\theta \\ &\overset{(6)}{\leq} X_1 + \varphi_1 \int_0^s M_3(|\bar{P}_1(\theta)| + |U_1(\theta)|)d\theta \\ &\leq X_1 + \varphi_1 M_3 \|U_1\|_{L^\infty(0,1)} + \varphi_1 M_3 \int_0^1 |\bar{P}_1(\theta)|d\theta \end{aligned} \tag{23}$$

Applying Gronwall's inequality yields

$$|\bar{P}_1(s)| \leq e^{\overline{D}M_3}(\overline{X} + M_3\overline{D}\,\overline{U})\,. \tag{24}$$

Then, applying (24) yields the following calculation

$$\begin{aligned} \bar{P}_1(s) - \bar{P}_2(s) &= X_1 - X_2 + \varphi_1 \int_0^s f(\bar{P}_1(\theta), U_1(\theta))d\theta \\ &\quad - \varphi_2 \int_0^s f(\bar{P}_2(\theta), U_2(\theta))d\theta \\ &\leq |X_1 - X_2| \\ &\quad + (\varphi_1 - \varphi_2) \int_0^s f(\bar{P}_1(\theta), U_1(\theta))d\theta \\ &\quad + \varphi_2 \int_0^s \Big(f(\bar{P}_1(\theta), U_1(\theta)) \\ &\quad - f(\bar{P}_2(\theta), U_2(\theta))\Big)d\theta \\ &\overset{(6)}{\leq} |X_1 - X_2| + (\varphi_1 - \varphi_2)M_3\|U_1\|_{L^\infty(0,s)} \\ &\quad + (\varphi_1 - \varphi_2)M_3 \int_0^s |P(\theta)|d\theta \\ &\quad + \varphi_2 M_1 \int_0^s |\bar{P}_1(\theta) - P_2(\theta)| \\ &\quad + |U_1(\theta) - U_2(\theta)|d\theta \\ &\overset{(24)}{\leq} |X_1 - X_2| \\ &\quad + |\varphi_1 - \varphi_2|M_3\left[\overline{U} + e^{M_3\overline{D}}(\overline{X} + M_3\overline{D}\,\overline{U})\right] \\ &\quad + \overline{D}M_3\|U_1 - U_2\|_{L^\infty(0,s)} \\ &\quad + \overline{D}M_3 \int_0^s |\bar{P}_1(\theta) - \bar{P}_2(\theta)|d\theta\,. \end{aligned} \tag{25}$$

Applying Gronwall's inequality again and letting $s = 1$ yields the desired result. ■

Given the continuity in Lemma 1, we can then apply Theorem 1 to the predictor operator yielding the existence of an arbitrarily close neural operator approximation:

Theorem 2 *Fix a compact set $K \subset \mathcal{X} \times C^2([0,1]; \mathcal{U}) \times \mathcal{D}$. Then, for all $\overline{X}$, $\overline{U}$, $\overline{D}$, $\epsilon > 0$, there exists a neural operator approximation $\breve{\mathcal{P}}_{NO} : K \to C^1([0,1]; \mathbb{R}^n)$ such* that for all $s \in [0,1]$,

$$\sup_{(X,U,\varphi)\in K} |\breve{\mathcal{P}}(X,U,\varphi)(s) - \breve{\mathcal{P}}_{NO}(X,U,\varphi)(s)| < \epsilon\,. \quad (26)$$

Now, we have established the existence of an arbitrarily close approximate predictor, but we did not discuss the detail on the number of parameters or data required to obtain this bound. Such analysis is beyond the goal of this paper, but we refer the reader to [44], [28], [36] for further details. Further, given the approximation is arbitrarily close in ϵ, this perturbation will affect the stability of the system compared to the global asymptotic result conducted with the exact predictor (which can never be implemented in practice). Therefore, in the next section, we identify the exact effect of the approximate predictor on the overall feedback loop.

4 Stability analysis under approximate predictors

We are now ready to analyze the plant (8a), (8b), (8c) with the neural operator approximated predictor

$$U(t) = \kappa(\breve{P}_{\mathrm{NO}}(t))\,, \quad (27)$$
$$\breve{P}_{\mathrm{NO}}(t) := \breve{\mathcal{P}}_{\mathrm{NO}}(X(t), T_D(t)U, \breve{D}(t))(1)\,, \quad (28)$$

where $T_D(t)$ is the historical control operator. The operator $T_D(t)$ shifts the function U back by $t - D(x-1)$ units such that, we have $u(x,t) = (T_D(t)U)(x)$ and therefore the resulting output of the exact predictor operator satisfies (12). Further, the choice of (28) is deliberate as it specifically operates on the function and vector directly - thereby avoiding a t dependence in the operator's domain. Notice that, in this case, we consider the full actuator measurement such that $u(x,t)$ is always known which is not guaranteed for every application. In Section 6, we extend this analysis to the unmeasured case.

We are now ready to present the main result of the paper.

Theorem 3 *Let the system* (8a), (8b), (8c) *satisfy Assumptions 1, 2, 3. Define the functional*

$$\Gamma(t) = |X(t)|^2 + \int_{t-D}^{t} U(\theta)^2 d\theta + |\tilde{D}(t)|^2\,, \quad (29)$$

where $\tilde{D}(t) = D - \breve{D}(t)$. Then, there exists parameters γ^*, $b^*(\bar{D}, M_3)$, $\overline{\Gamma}(\overline{X}, \overline{U}, \overline{D}) > 0$, a class $\mathcal{KL}$function β_1^* and class $\mathcal{K}_\infty$ functions $\alpha_{1:5}^*$such that if $\gamma < \gamma^*$, $b > b^*$, $\epsilon < \epsilon^*$ where

$$\epsilon^* = (\alpha_1^*)^{-1}(\overline{\Gamma} - \alpha_2^*(\overline{\Delta D}))\,, \quad (30)$$

where $\overline{\Delta D} = \overline{D} - \underline{D}$ and the initial state is constrained to

$$\Gamma(0) \le \alpha_3^*(\overline{\Gamma} - \alpha_1^*(\epsilon) - \alpha_2^*(\overline{\Delta D}))\,, \quad (31)$$

then,

$$\Gamma(t) \le \alpha_3^*(\Gamma(0)) + \alpha_1^*(\epsilon) + \alpha_2^*(\overline{\Delta D})\,, \quad (32)$$

and

$$|X(t)|^2 \le \beta_1^*(\Gamma(0), t) + \alpha_4^*(\overline{\Delta D}) + \alpha_5^*(\epsilon)\,, \quad (33)$$
$$\|u(\cdot,t)\|^2_{L^2(0,1)} \le \beta_1^*(\Gamma(0), t) + \alpha_4^*(\overline{\Delta D}) + \alpha_5^*(\epsilon)\,. \quad (34)$$

Notice that Theorem 3 is weaker than the result in [10, Theorem 7] because it depends on the delay projection bounds $\overline{D}$, $\underline{D}$. This is expected since the error introduced by the predictor is additive, leading to complications similar to those in robust adaptive control, where regulation depends on projector bounds (See [20]). Additionally, it may seem counterintuitive that ϵ^* increases with $\overline{\Gamma}$, but this is expected: a larger radius of states allows for a larger transient, thus loosening the ϵ^* bound.

Finally, we compare Theorem 3 with related adaptive neural-operator control results. In [11, Theorem 4] and [33, Theorem 4], the gain kernel is approximated, introducing a multiplicative error and yielding global stability, but requiring approximation of the time derivative of the update parameter. By contrast, Theorem 3 approximates $\breve{P}$, analogous to the direct control law in [10, Theorem 3]. This leads to an additive error and a semi-global practical result without requiring the derivative. Hence, the choice of approximation is important yielding a trade-off between global and semi-global guarentees.

PROOF. We first introduce a backstepping transform, that under the feedback law (27), transforms the system into a target system for the stability analysis.

Lemma 2 *Under the backstepping transformation with the* exact predictor *given by*

$$w(x,t) = u(x,t) - \kappa(\breve{p}(x,t))\,, \quad (35)$$

the system (8a), (8b), (8c) *with control law given by* (27) *is transformed into*

$$\dot{X}(t) = f(X(t), \kappa(X(t)) + w(0,t))\,, \quad (36a)$$
$$Dw_t = w_x - \tilde{D}(t)q_1(x,t) - D\dot{\breve{D}}(t)q_2(x,t)\,, \quad (36b)$$
$$w(1,t) = \kappa(\breve{p}_{NO}(1,t)) - \kappa(\breve{p}(1,t))\,, \quad (36c)$$

where $\tilde{D} = D - \breve{D}(t)$, $\breve{p}_{NO}(x,t) = \breve{\mathcal{P}}_{NO}(X(t), T_D(t)U, \hat{D}(t))(x)$, and q_2 is given by

$$q_2 = \frac{\partial \kappa}{\partial \breve{p}}(\breve{p}(x,t)) \int_0^x \Phi(x,y,t) f(\breve{p}(y,t), \kappa(\breve{p}(y,t)) + w(y,t))dy\,. \quad (37)$$

Because we use the exact backstepping transform in (35), the proof of Lemma 2 follows that of [12, Lemma 1], except for the boundary term at $x = 1$, which is obtained by directly substituting $U(t) = \kappa(\breve{p}_{\mathrm{NO}}(\cdot))$ into (35).

To analyze (36a), (36b), (36c), we introduce the following Lyapunov-Krasovskii functional

$$W(t) = D \log N(t) + \frac{b}{\gamma}\tilde{D}(t)^2 \quad (38)$$

$$N(t) = 1 + V(x) + b\int_0^1 (1+x)w(x,t)^2 dx\,. \quad (39)$$

Taking the time derivative and substituting w_t yields

$$\begin{aligned}\dot{W}(t) =& \frac{1}{N(t)}\bigg(D\frac{\partial V(X)}{\partial X}\Big[f(X(t),\kappa(X(t))+w(0,t))\Big]\bigg)\\ &+\frac{2b}{N(t)}\int_0^1 (1+x)w(x,t)\bigg(w_x(x,t)-\tilde{D}(t)q_1(x,t)\\ &-D\dot{\hat{D}}(t)q_2(x,t)\bigg)dx - \frac{2b}{\gamma}\tilde{D}\dot{\hat{D}}(t)\,. \quad (40)\end{aligned}$$

Using the Lipschitzness of f in Assumption 3 and the Lyapunov condition in Assumption 2, we have

$$\begin{aligned}\dot{W}(t) \le& \frac{1}{N(t)}\left(-D\lambda|X(t)|^2 + DC_2M_3|X(t)||w(0,t)|\right)\\ &+\frac{2b}{N(t)}\int_0^1 (1+x)w(x,t)\bigg(w_x(x,t)-\tilde{D}(t)q_1(x,t)\\ &-D\dot{\hat{D}}(t)q_2(x,t)\bigg)dx - \frac{2b}{\gamma}\tilde{D}\dot{\hat{D}}\,. \quad (41)\end{aligned}$$

Distributing the second term, applying integration by parts, and substituting (14) yields

$$\begin{aligned}\dot{W}(t) \le& \frac{1}{N(t)}\bigg(-D\lambda|X(t)|^2 + DC_2M_1|X(t)||w(0,t)|\\ &+bw^2(1,t) - bw^2(0,t) - b\|w(\cdot,t)\|^2_{L^2(0,1)}\\ &+2bD\dot{\hat{D}}(t)\int_0^1 (1+x)w(x,t)q_2(x,t)dx\bigg)\\ &+\frac{2b}{\gamma}\tilde{D}(t)(\gamma\phi(t)-\dot{\hat{D}}(t))\,. \quad (42)\end{aligned}$$

Using [12, Lemma 2 and 3] which states that there exists constants $M_6, M_6, M_7 > 0$ such that

$$M_6(|X| + \|w(\cdot,t)\|_2) \ge |\breve{p}(x,t)|\,, \quad (43)$$

$$M_6\left(|X|^2 + \|w(\cdot,t)\|_2^2\right) \ge 2bD\left|\int_0^1 (1+x) \times q_2(x,t)w(x,t)dx\right|, \quad (44)$$

$$\gamma M_7 \ge \left|\dot{\hat{D}}(t)\right|\,, \quad (45)$$

in conjunction with Young's inequality yields

$$\begin{aligned}\dot{W}(t) \le& \frac{1}{N(t)}\bigg(-\eta\left(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)}\right)\\ &-\left(b - \frac{C_2^2DM_3^2}{2\lambda}\right)w^2(0,t)\\ &+\gamma M_6M_7\left(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)}\right)\\ &+bw^2(1,t)\bigg)\,, \quad (46)\end{aligned}$$

where $\eta := \min(D\lambda/2, b)$. Thus, given that

$$b > \frac{C_2^2\overline{D}M_3^2}{2\lambda} =: b^*\,, \quad (47)$$

$$\gamma < \frac{\eta}{M_6M_7} =: \gamma^*\,, \quad (48)$$

guarantees there exists $C > 0$ such that

$$\begin{aligned}\dot{W}(t) \le& -\frac{C}{N(t)}(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)})\\ &+\frac{b}{N(t)}(\kappa(\breve{p}(1,t)) - \kappa(\breve{p}_{\mathrm{NO}}(1,t)))^2\,. \quad (49)\end{aligned}$$

Now, using the definition of $W(t)$ and Assumption 2, we have that

$$\begin{aligned}e^{W(t)/D - \frac{b}{\gamma D}(\tilde{D}(t))^2} - 1 \le& \max\{C_1, 2b\}\\ &\times(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)})\,,\end{aligned} \quad (50)$$

Using reverse Young's inequality (See [8, Appendix A]), we have

$$e^{W(t)/D}e^{-\frac{b}{\gamma D}(\tilde{D}(t))^2} \ge 2e^{W(t)/(2D)} - e^{\frac{b}{\gamma D}\tilde{D}(t))^2} \quad (51)$$

yielding

$$\begin{aligned}2e^{W(t)/D} - &e^{\frac{b}{\gamma D}(\tilde{D}(t))^2} - 1\\ &\le \max\{C_1, 2b\}(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)})\,,\end{aligned} \quad (52)$$

which then results in

$$\begin{aligned}-(|X(t)|^2 + \|w(\cdot,t)\|^2_{L^2(0,1)}) \le& \frac{1}{\max\{C_1, 2b\}}\bigg(-2e^{W(t)/D}\\ &+e^{\frac{b}{\gamma D}(\tilde{D}(t))^2} + 1\bigg)\,.\end{aligned} \quad (53)$$

Substituting into (49) yields

$$\begin{aligned}\dot{W}(t) \le& -\frac{2C}{\max\{C_1,2b\}N(t)}e^{\frac{W(t)}{D}}\\ &+\bigg[\frac{C}{\max\{C_1,2b\}N(t)}\left(1 + e^{\frac{b}{D\gamma}(\overline{\Delta D})^2}\right)\\ &+b(\kappa(\breve{p}(1,t)) - \kappa(\breve{p}_{\mathrm{NO}}(1,t)))^2\bigg]\\ \le& -\frac{2C}{\max\{C_1,2b\}N(t)}\left(e^{\frac{W(t)}{D}} - 1\right)\\ &+\bigg[\frac{C}{\max\{C_1,2b\}N(t)}\left(e^{\frac{b}{D\gamma}(\overline{\Delta D})^2} - 1\right)\\ &+b(\kappa(\breve{p}(1,t)) - \kappa(\breve{p}_{\mathrm{NO}}(1,t)))^2\bigg]\,. \quad (54)\end{aligned}$$

Following [32, Theorem C.3], we have that there exists functions $\beta_1 \in \mathcal{KL}$ and $\alpha_{1:2} \in \mathcal{K}_\infty$ such that

$$\begin{aligned} W(t) \leq& \beta_1(W(0), t) + \alpha_1(\overline{\Delta D}) \\ &+ \alpha_2(\sup_{0\leq\tau\leq t} (\kappa(\breve{p}(1,\tau)) - \kappa(\breve{p}_{\rm NO}(1,\tau)))^2) . \end{aligned} \quad (55)$$

Now, let us define the sub-level set bounded by $\overline{W}$ such that $\Omega_{\overline{W}} := \{z : W(z) \leq \overline{W}\}$. Then, on the set of trajectories in $\Omega_{\overline{W}}$, we can apply Theorem 2 such that

$$\begin{aligned} &(\kappa(\breve{p}(1,t)) - \kappa(\breve{p}_{\rm NO}(1,t)))^2 \\ &\quad\leq M_1^2|\breve{p}(1,t) - \breve{p}_{\rm NO}(1,t)|^2 \\ &\quad\leq M_1^2|\breve{\mathcal{P}}(X(t), T_D(t)U, \breve{D}(t))(1) \\ &\qquad - \breve{\mathcal{P}}_{\rm NO}(X(t), T_D(t)U, \breve{D}(t))(1)|^2 \\ &\quad\leq M_1^2 \sup_{x\in[0,1]} |\breve{\mathcal{P}}(X(t), T_D(t)U, \breve{D}(t))(x) \\ &\qquad - \breve{\mathcal{P}}_{\rm NO}(X(t), T_D(t)U, \breve{D}(t))(x)|^2 , \\ &\quad\leq M_1^2\epsilon^2 . \end{aligned} \quad (56)$$

Hence, we obtain $\alpha_3 \in \mathcal{K}_\infty$ such that

$$W(t) \leq \beta_1(W(0), t) + \alpha_1(\overline{\Delta D}) + \alpha_3(\epsilon) . \quad (57)$$

To stay in the compact set $\Omega_{\overline{W}}$, we need to impose conditions on $W(0)$ and ϵ^*. We first convert back to the full function Γ and then impose the corresponding conditions. Now, (43) and Assumption's 2, 3 imply that there exists constants $r_1, r_2, s_1, s_2 > 0$ such that

$$\|u(\cdot,t)\|^2_{L^2(0,1)} \leq r_1|X(t)|^2 + r_2\|w(\cdot,t)\|^2_{L^2(0,1)} , \quad (58)$$
$$\|w(\cdot,t)\|^2_{L^2(0,1)} \leq s_1|X(t)|^2 + s_2\|u(\cdot,t)\|^2_{L^2(0,1)} . \quad (59)$$

To obtain an estimate on Γ, note that from [12, Eqn. (47), (48)], we have that

$$\Gamma(t) \leq \left(D\left(r_1 + \frac{r_2}{b}\right) + 1 + \frac{\gamma D}{b}\right)\left(e^{\frac{W(t)}{D}} - 1\right) , \quad (60)$$
$$W(t) \leq D\left(C_1 + 2b\left(s_1 + \frac{s_2}{D}\right) + \frac{b}{\gamma D}\right)\Gamma(t) . \quad (61)$$

Therefore, using the stability estimate in (57), we obtain the existence of class K^∞ functions α_1^*, α_2^*, α_3^* such that

$$\Gamma(t) \leq \alpha_3^*(\Gamma(0)) + \alpha_1^*(\epsilon) + \alpha_2^*(\overline{\Delta D}) , \quad (62)$$

when

$$\epsilon < \epsilon^* := (\alpha_1^*)^{-1}(\overline{\Gamma} - \alpha_2^*(\overline{\Delta D})) , \quad (63)$$
$$\Gamma(0) \leq (\alpha_3^*)^{-1}(\overline{\Gamma} - \alpha_1^*(\epsilon) - \alpha_2^*(\overline{\Delta D})) \quad (64)$$

where $\overline{\Gamma} = \overline{X} + \overline{U} + \overline{D}$. Note that the conditions (63), (64) ensure that the set $\Omega_{\overline{W}}$ is forward invariant ensuring that Theorem 2 is always applied on a compact set.

To show obtain a bound on $X(t)$, note that we have

$$|X(t)|^2 \leq V(t) \leq (e^{\frac{W(t)}{D}} - 1) . \quad (65)$$

Using the stability estimate in (57) and inequalities (65), there exists $\beta_2 \in \mathcal{KL}$, $\alpha_6, \alpha_7 \in \mathcal{K}_\infty$ such that

$$|X(t)|^2 \leq \beta_2(W(0), t) + \alpha_6(\overline{\Delta D}) + \alpha_7(\epsilon) . \quad (66)$$

Similarly, to show a stability estimate on $U(t)$, we use (58), (59) with the definition of $W(t)$, yielding

$$\begin{aligned} \int_{t-D}^t U(s)^2 ds &= D\|u(\cdot,t)\|^2_{L^2(0,1)} \\ &\leq D(r_1|X(t)|^2 + r_2\|w(\cdot,t)\|^2_{L^2(0,1)}) \\ &\leq D\left(r_1 + \frac{r_2}{b}\right)\left(e^{\frac{W(t)}{D}} - 1\right) . \end{aligned} \quad (67)$$

Then, repeating the argument above with the estimate in (57), there exists $\beta_3 \in \mathcal{KL}$, $\alpha_8, \alpha_9 \in \mathcal{K}_\infty$ such that

$$\|u(t)\|^2_{L^2(0,1)} \leq \beta_3(W(0), t) + \alpha_8(\overline{\Delta D}) + \alpha_9(\epsilon) . \quad (68)$$

Letting $\beta_1^*(\cdot,\cdot) = \max\{\beta_2(\cdot,\cdot), \beta_3(\cdot,\cdot)\}$, $\alpha_4^*(\cdot) = \max\{\alpha_6(\cdot), \alpha_8(\cdot)\}$ and $\alpha_5^*(\cdot) = \max\{\alpha_7(\cdot), \alpha_9(\cdot)\}$ and applying the inequality (60) completes the result. ■

5 Numerical Experiments

All code, parameters, and datasets are available [41].

Consider the biological system in [15] consisting of two proteins: an activator protein that promotes expression of itself and a repressor protein that represses the expression of the activator. Such biological clocks play fundamental roles in cell physiology [47] and control of such systems is of valuable interest to synthetic biologists for the design of new medicines. The dyanmics are given by

$$\dot{x}_1 = -x_1 + f_1(x_1, x_2) + U(t - D) , \quad (69)$$
$$\dot{x}_2 = -\frac{x_2}{2} + f_2(x_1) , \quad (70)$$

where x_1 is the concentration of the activator protein, x_2 is the concentration of the repressor protein and f_1, f_2 are Hill functions given by

$$f_1(x_1, x_2) = \frac{K_1x_1^2 + K_a}{1 + x_1^2 + x_2^2} , \quad (71)$$
$$f_2(x_1) = \frac{K_2x_1^2 + K_b}{1 + x_1^2} , \quad (72)$$

with $K_1 = K_2 = 300, K_a = 0.04, K_b = 0.004$. The input U controls the activator protein concentration with a delay due to the time to pass through an inlet pipe. Additionally, we assume the pipe is equipped with a UV spectrometer such that we can measure the presence of the activator protein along the pipe and thus the distributed input $u(x,t) = U(t - D(1 - x))$ is known.

A control law satisfying Assumption 2 is given by

$$\kappa(X(t)) = -f_1(x_1, x_2) + f_1(x_1^*, x_2^*) , \quad (73)$$

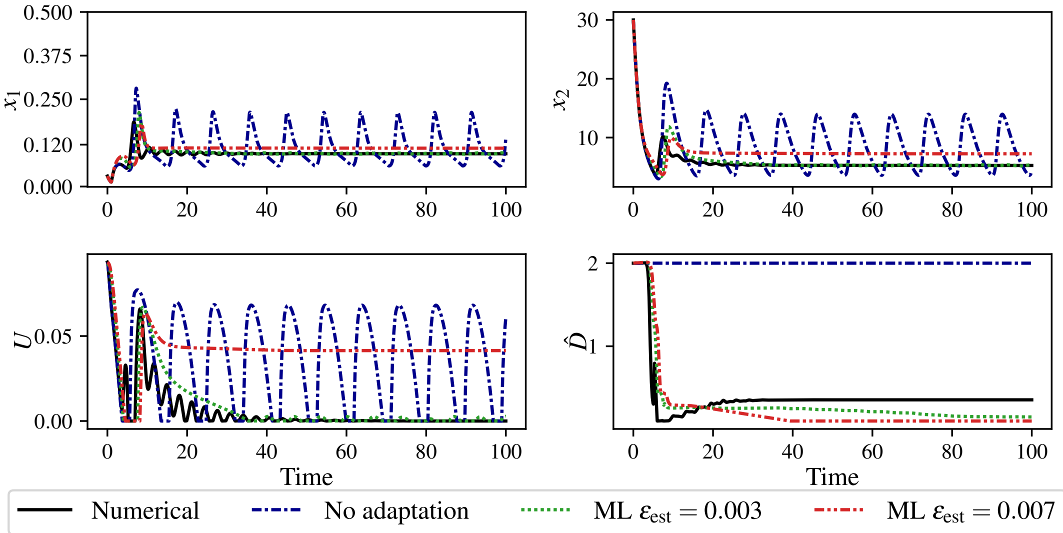

Fig. 1. Simulation of the plant (1) with various approximate predictors. The initial state is $X(0) = [0.03, 30]$, delay is $D = 1$, $\hat{D}(0) = 2$, $\gamma = 1000$, $b = 1$ which is within the training distribution, but not explicitly seen. The black line indicates the numerical predictor, the blue-line is the predictor without delay-adaptation, the red line indicates the DeepONet predictor with higher error and the green line indicates the DeepONet predictor trained to optimality.

where $(x_1^*, x_2^*) = (0.0939, 5.2525)$ is an unstable equilibrium setpoint of the system. The associated Lypaunov function is given by $V = (X - X^*)^T (X - X^*)$ and note the system experiences a limit cycle [12, Figure 2] when an open-loop control law $U(\cdot) = 0$ is used.

To train the neural operator, we generate a dataset of 5000 instances by simulating the system with a numerical predictor [24] across various initial conditions and delay estimates (see code for parameters; dataset generation took 60 minutes). Two neural operator architectures (DeepONet, FNO) are trained in 10 minutes on an NVIDIA 3090Ti GPU. Both perform similarly, though DeepONet achieves a larger speedup (Table 1).

Figure 1 shows system simulations using the DeepONet predictor. To illustrate the effect of ϵ, we compare a predictor trained with early stopping (large ϵ) and one trained to optimality: as shown by the red line, early stopping yields convergence to a larger neighborhood of the equilibrium, while even the optimally trained predictor does not exactly reach the setpoint, converging to $(0.094, 5.39)$ rather than $(0.0939, 5.2525)$ due to the uniform approximation error ϵ. In both cases, ϵ determines the convergence radius, consistent with Theorem 3.

Finally, we evaluated the computational speedup of the neural operator predictor. Across all discretizations, the ML predictor outperforms the numerical approach, achieving up to $15\times$ speedup at $dx = 0.001$, with stability requiring $dx < 0.005$ (Euler scheme). We note this speedup improvement is lower than the $100\times$ reported in [9, Table 1] due to the simpler dynamics in this example; however, applying the same approach to more computationally expensive forward dynamics would likely yield even greater gains with DeepONet.

Table 1
Computation time for approximate predictors averaged over 1000 samples for the biological protein example (seconds).

Step size (dx)	Numerical ↓	DeepONet ↓	FNO ↓	DeepONet Speedup ↑	FNO Speedup ↑
0.01	1.601	**0.496**	1.331	**3.22x**	1.20x
0.005	3.295	**0.587**	1.440	**5.61x**	2.29x
0.001	18.197	**1.212**	2.108	**15.01x**	8.63x

6 Control in presence of an unmeasured distributed input

In Section 4, we assumed the actuator state $u(x, t)$ was known. In many real-world systems, the actuator is unknown and subject to uncertain or unmeasured delays. Here, we extend Section 4 to unknown actuator states. We introduce an **estimate of the actuator state** as

$$\hat{u}(x, t) = U(t + \hat{D}(t)(x - 1)), \quad x \in [0, 1], \tag{74}$$

where $\hat{D}$ is an estimate of the delay. The control law is now computed in the same approach as before, but with the estimated state $\hat{u}$:

$$U(t) = \kappa(\hat{p}(1, t)), \tag{75}$$

$$\begin{aligned} \hat{p}(x, t) &= X(t + \hat{D}(t)x) \\ &= X(t) + \hat{D}(t) \int_0^x f(\hat{p}(y, t), \hat{u}(y, t)) dy, \end{aligned} \tag{76}$$

where the predictor $\hat{p}$ is computed using the approximate actuator state $\hat{u}(x, t)$ as opposed to $u(x, t)$ in (12). Hence, the neural operator approximated feedback law becomes

$$U(t) = \kappa(\hat{p}_{\text{NO}}(1, t)), \tag{77}$$

$$\hat{p}_{\text{NO}}(x, t) := \breve{\mathcal{P}}_{\text{NO}}(X(t), T_{\hat{D}}(t)U, \hat{D}(t))(x), \tag{78}$$

where the main difference is that the distributed input is now given by $T_{\hat{D}}(t)U$ using (74).

Notice that, due to the requirement of the estimated distributed state $\hat{u}$, the analysis will become significantly more complex then Section 4 as one needs to study the ODE-PDE-PDE cascade (c.f. Lemma 3) containing not only X, u, and $\tilde{u} = u - \hat{u}$, but also their corresponding spatial and time derivatives. This is more intricate then just ODE-PDE cascade where $\tilde{D}$ appears linearly only in the PDE estimate (36b). Hence, developing a stability result in even the exact predictor case will require the update law belongs to a class of functions of which can be regulated by the state and derivatives of the actuation:

Assumption 4 *There exists a C^1 function $\hat{\phi}$, a positive parameter $\gamma > 0$ and class $\mathcal{K}$ functions ρ_1, ρ_2 such that*

$$\dot{\hat{D}}(t) = \gamma Proj_{[\underline{D}, \overline{D}]}\left\{\hat{D}(t)\,, \hat{\phi}(t)\right\}\,, \tag{79}$$

where $\hat{\phi}(t)$ satisfies

$$|\hat{\phi}(t)| \leq \rho_1(\Xi(t)) \text{ and } |\dot{\hat{\phi}}(t)| \leq \rho_2(\Xi(t)) \tag{80}$$

with the function $\Xi(t)$ given by

$$\begin{aligned}\Xi(t) = |X(t)| &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} |U(s)| ds \\ &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} |\dot{U}(s)| ds \\ &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} \left|\ddot{U}(s)\right| ds\,.\end{aligned} \tag{81}$$

Notice that Assumption 4 is strong, but there are a series of estimators that satisfy this Assumption. For example, we will consider the design as in [12] governed by

$$\begin{aligned}\hat{\phi}(t) \in\, & 2\,\mathrm{Sgn}\big(\hat{w}_x(1,t)\big)\, q_3(1,t) \\ &+ \int_0^1 (1+x)\Big[q_3(x,t)\,\mathrm{Sgn}\big(\hat{w}(x,t)\big) \\ &\qquad + q_4(x,t)\,\mathrm{Sgn}\big(\hat{w}_x(x,t)\big)\Big] dx,\end{aligned} \tag{82}$$

where

$$\mathrm{Sgn}(r) = \begin{cases} \{1\}, & r > 0, \\ [-1,1], & r = 0, \\ \{-1\}, & r < 0\,, \end{cases}$$

where update law is interpreted in the Filippov sense as a differential inclusion [18] and $\hat{w}$ is given by

$$\hat{w}(x,t) = \hat{u}(x,t) - \kappa(\hat{p}(x,t))\,, \tag{83}$$

and the functions q_3, q_4 are given by

$$q_3(x,t) = \frac{d\kappa}{d\hat{p}}(\hat{p}(x,t)) f(\hat{p}(0,t), \hat{u}(0,t))\,, \tag{84}$$

$$q_4(x,t) = \frac{\partial}{\partial x} q_3(x,t)\,. \tag{85}$$

This update law is guided by the same Lyapunov–Krasovskii intuition as in Section 4. The law can be viewed as an L_1-type steepest-descent mechanism built from $\hat{w}$ and $\hat{w}_x$ to compensate the unfavorable terms in the Lyapunov derivative.

We are now ready to state the our main result.

Theorem 4 *Consider the plant* (8a), (8b), (8c) *satisfying Assumptions 1, 2, 3, 4 and the functional*

$$\begin{aligned}\Upsilon(t) = |X(t)| &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} |U(s)| ds \\ &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} |\dot{U}(s)| ds \\ &+ \int_{t-\max\{D,\hat{D}(t)\}}^{t} |\ddot{U}(s)| ds + \tilde{D}(t)^2\,.\end{aligned} \tag{86}$$

Then, there exists functions $\beta_2^* \in \mathcal{KL}$ and $\alpha_6^*, \alpha_7^*, \alpha_8^*, \alpha_9^* \in \mathcal{K}_\infty$ and constants $\overline{\Upsilon} > 0$, $\gamma^*(\Upsilon(0)) > 0$,

$$\hat{\epsilon}^*(\overline{\Upsilon}) := (\alpha_6^*)^{-1}(\overline{\Upsilon}) > 0\,, \tag{87}$$

and a region of attraction

$$R_\Upsilon(\overline{\Upsilon}, \epsilon) := \alpha_7^*(\overline{\Upsilon} - \alpha_6^*(\epsilon)) > 0\,, \tag{88}$$

such that if $\gamma < \gamma^*$, $\epsilon < \hat{\epsilon}^*$, and the initial state is constrained to $|\Upsilon(0)| \leq R_\Upsilon$, then, the feedback law (77) *with the estimator* (79), (82) *guarantees for all solutions*

$$\Upsilon(t) \leq \alpha_7^*(\Upsilon(0)) + \alpha_6^*(\epsilon)\,, \quad \forall t > 0\,, \tag{89}$$

and

$$\begin{aligned}|X(t)| &\leq \beta_2^*(\Upsilon(0), t) + \alpha_8^*(\overline{\Delta D}) + \alpha_9^*(\epsilon)\,, &(90)\\ \|u(\cdot,t)\|_{L^1} &\leq \beta_2^*(\Upsilon(0), t) + \alpha_8^*(\overline{\Delta D}) + \alpha_9^*(\epsilon)\,. &(91)\end{aligned}$$

Theorem 4 is inherently a local result in that it requires the set of initial condition radius to be smaller then R_Υ which inherently decreases as the approximation error worsens. Hence, only for $\epsilon < \epsilon^*$ does a local region of attraction exist, but qualitatively, the result ensures that an ϵ^* always exists - although it can be extremely small and inherently unachievable in practice. In the case of which $\epsilon = 0$, we recover the local result of [12, Theorem 3]. Furthermore, it may seem odd that the analysis and hence the result is characterized in the L^1 norm; however, as we will see the coupling between the estimation error and the stability of the target system will be in an absolute sense. Hence, one would require additional assumptions for the L^2 norm (c.f. [12]).

PROOF. The proof has three parts: derivation of the target PDE–PDE–ODE cascade with the actuated measurement, construction and differentiation of a Lyapunov–Krasovskii functional using Lemmas 4–10, and conditions ensuring forward invariance of the Lyapunov functional, which establish the local result.

Deriving the target system First, we begin by characterizing the PDE describing $\hat{u}$ as well as the corresponding target system. By direct calculation, we have

$$\frac{\partial \hat{u}(x,t)}{\partial x} = U'(t+\hat{D}(t)(x-1))\hat{D}(t)\,, \tag{92}$$

$$\frac{\partial \hat{u}(x,t)}{\partial t} = U'(t+\hat{D}(t)(x-1))(1+\dot{\hat{D}}(t)(x-1))\,. \tag{93}$$

Therefore, $\hat{u}(x,t)$ satisfies the PDE

$$\hat{D}(t)\hat{u}_t(x,t) = \hat{u}_x(x,t) + \dot{\hat{D}}(t)(x-1)\hat{u}_x(x,t)\,, \tag{94a}$$

$$\hat{u}(1,t) = U(t)\,. \tag{94b}$$

Thus, under the backstepping transform with the exact predictor, we obtain the following target system:

Lemma 3 *Under the backstepping transformation for the distributed system* (94a), (94b) *given by*

$$\hat{w}(x,t) = \hat{u}(x,t) - \kappa(\hat{p}(x,t)), \tag{95}$$

the plant (1) *becomes*

$$\dot{X}(t) = f(X(t), \kappa(X(t)) + \hat{w}(0,t) + \hat{u}(0,t)), \tag{96a}$$

$$\begin{aligned}\hat{D}(t)\hat{w}_t(x,t) =& \hat{w}_x(x,t) + \dot{\hat{D}}(t)g_3(x,t)\\ &- g_4(x,t)f_{\tilde{u}}(t)\end{aligned} \tag{96b}$$

$$\hat{w}(1,t) = \kappa(\hat{p}_{NO}(1,t)) - \kappa(\hat{p}(1,t))\,, \tag{96c}$$

$$\begin{aligned}D\tilde{u}_t(x,t) =& \tilde{u}_x(x,t) - \tilde{D}(t)g_1(x,t)\\ &- \dot{\hat{D}}(t)g_2(x,t)\,,\end{aligned} \tag{96d}$$

$$\tilde{u}(1,t) = 0 \tag{96e}$$

where $\tilde{u} = u - \hat{u}$ represents the difference between the true actuation and the estimate.

$$\begin{aligned}g_1(x,t) :=& \frac{1}{\hat{D}(t)}\left[\hat{w}_x(x,t) + \hat{D}(t)\frac{d\kappa(\hat{p}(x,t))}{d\hat{p}}\right.\\ &\left.\times f(\hat{p}(x,t), \hat{w}(x,t) + \kappa(\hat{p}(x,t)))\right]\,,\end{aligned} \tag{97}$$

$$\begin{aligned}g_2(x,t) :=& \frac{D(x-1)}{\hat{D}(t)}\left[\hat{w}_x(x,t) + \hat{D}(t)\frac{d\kappa(\hat{p}(x,t))}{d\hat{p}}\right.\\ &\left.\times f(\hat{p}(x,t), \hat{w}(x,t) + \kappa(\hat{p}(x,t)))\right]\,,\end{aligned} \tag{98}$$

$$\begin{aligned}g_3(x,t) :=& (x-1)\Big[\hat{w}_x(x,t) + \kappa'(\hat{p}(x,t))\hat{D}(t)\\ &\times f(\hat{p}(x,t), \hat{w}(x,t) + \kappa(\hat{p}(x,t)))\Big]\\ &- \kappa'(\hat{p}(x,t))\hat{D}(t)\int_0^x \hat{\Phi}(x,y,t)\\ &\times \Big[f(\hat{p}(y,t), \hat{w}(y,t) + \kappa(\hat{p}(y,t)))\\ &+ \frac{\partial f(\hat{p}(y,t), \hat{w}(y,t) + \kappa(\hat{p}(y,t)))}{\partial \hat{u}}(y-1)\\ &\times \Big(\hat{w}_x(y,t) + \kappa'(\hat{p}(y,t))\hat{D}(t)f(\hat{p}(y,t), \hat{w}(y,t)\\ &+ \kappa(\hat{p}(y,t)))\Big)\Big]dy\end{aligned} \tag{99}$$

$$g_4(x,t) := \hat{\Phi}(x,0,t)\hat{D}(t)\kappa'(\hat{p}(x,t))\,, \tag{100}$$

$$f_{\tilde{u}}(t) := f(\hat{p}(0,t), u(0,t)) - f(\hat{p}(0,t), \hat{u}(0,t))\,, \tag{101}$$

and $\hat{\Phi}$ is the state transition matrix associated with the spatially-varying differential equation governing $\hat{r}(x) = \hat{D}(t)\hat{p}_t(x,t) - \hat{p}_x(x,t)$.

See [8, Appendix A] for proof.

Note, as in Lemma 2, the challenge of the target system is the neural operator error in (96c). By direct calculation of (96d), (96e), (96a), (96b):

$$\begin{aligned}D\tilde{u}_{xt}(x,t) =& \tilde{u}_{xx}(x,t) - \tilde{D}(t)g_5(x,t)\\ &- \dot{\hat{D}}(t)g_6(x,t)\,,\end{aligned} \tag{102}$$

$$\tilde{u}_x(1,t) = \tilde{D}(t)g_1(1,t)\,, \tag{103}$$

$$\begin{aligned}\hat{D}(t)\hat{w}_{xt}(x,t) =& \hat{w}_{xx}(x,t) + \dot{\hat{D}}(t)g_7(x,t)\\ &- g_8(x,t)f_{\tilde{u}}(t)\,,\end{aligned} \tag{104}$$

$$\hat{w}_x(1,t) = -\dot{\hat{D}}(t)g_3(1,t) + g_4(1,t)f_{\tilde{u}}(t)\,, \tag{105}$$

$$\begin{aligned}\hat{D}(t)\hat{w}_{xxt}(x,t) =& \hat{w}_{xxx}(x,t) + \dot{\hat{D}}(t)g_9(x,t)\\ &- g_{10}(x,t)f_{\tilde{u}}(t)\,,\end{aligned} \tag{106}$$

$$\begin{aligned}\hat{w}_{xx}(1,t) =& -\dot{\hat{D}}(t)g_7(1,t) + g_8(1,t)f_{\tilde{u}}(t)\\ &+ \hat{D}(t)\hat{w}_{xt}(1,t)\,,\end{aligned} \tag{107}$$

where $g_5 = g_{1,x}$, $g_6 = g_{2,x}$, $q_7 = g_{3,x}$, $q_8 = q_{4,x}$, $q_9 = q_{7,x}$, and $g_{10} = g_{8,x}$. We refer the reader to [8, Appendix C] for the full expressions and omit them for brevity here.

Derivative of the Lyapunov-Krasovskii functional Consider now the Lyapunov-Krasovskii candidate

$$\begin{aligned}\hat{W}(t) =& V_0(X) + b_0 D\int_0^1 (1+x)|\tilde{u}(x,t)|dx\\ &+ b_1 D\int_0^1 (1+x)|\tilde{u}_x(x,t)|dx\\ &+ b_2\hat{D}(t)\int_0^1 (1+x)|\hat{w}(x,t)|dx\\ &+ b_3\hat{D}(t)\int_0^1 (1+x)|\hat{w}_x(x,t)|dx\\ &+ b_4\hat{D}(t)\int_0^1 (1+x)|\hat{w}_{xx}(x,t)|dx + \tilde{D}(t)^2\,,\end{aligned} \tag{108}$$

where $V_0 = \sqrt{V}$ as in Assumption 2 and $b_{0:4} > 0$ are constants to be specified. By the mean value theorem:

$$\dot{V}_0 = \frac{1}{2}\frac{dV_0}{dX}(f(X, \kappa(X) + \tilde{u}(0,t) + \hat{w}(0,t))$$

$$\leq -\frac{\lambda}{2}|X| + \frac{C_2}{2}\left|\frac{\partial f(X, a_1(X,t))}{\partial u}\right| |\hat{w}(0,t) + \tilde{u}(0,t)| \tag{109}$$

where $a_1(X,t) \in [\kappa(X(t)), \kappa(X(t)) + \hat{u}(0,t)]$. We begin by differentiating with respect to time and substitute the corresponding PDEs for $\tilde{u}_t$, $\tilde{u}_{xt}$, $\hat{w}_t$, $\hat{w}_{xt}$. The resulting integrals will be given in terms on only spatial derivatives and hence we can apply integration by parts yielding

$$\begin{aligned}
\dot{\hat{W}}(t) \leq & -\lambda/2|X| + \frac{C_2}{2}\left|\frac{\partial f(x, \alpha_1(x,t))}{\partial u}\right| |\hat{w}(0,t) + \tilde{u}(0,t)| \\
& + b_0\Big[-|\tilde{u}(0,t)| - \|\tilde{u}(\cdot,t)\|_{L^1} \\
& + |\tilde{D}(t)| \int_0^1 (1+x)|g_1(x,t)|dx \\
& + |\dot{\hat{D}}(t)| \int_0^1 (1+x)|g_2(x,t)|dx \Big] \\
& + b_1\Big[2|\tilde{u}_x(1,t)| - |\tilde{u}_x(0,t)| - \|\tilde{u}_x(\cdot,t)\|_{L^1} \\
& + |\tilde{D}(t)| \int_0^1 (1+x)|g_5(x,t)|dx \\
& + |\dot{\hat{D}}(t)| \int_0^1 (1+x)|g_6(x,t)|dx \Big] \\
& + b_2\Big[2|\hat{w}(1,t)| - |\hat{w}(0,t)| - \|\hat{w}(\cdot,t)\|_{L^1} \\
& + |\dot{\hat{D}}(t)| \int_0^1 (1+x)|g_3(x,t)|dx \\
& |f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_4(x,t)|dx \Big] \\
& + b_3\Big[2|\hat{w}_x(1,t)| - |\hat{w}_x(0,t)| - \|\hat{w}_x(\cdot,t)\|_{L^1} \\
& + |\dot{\hat{D}}(t)| \int_0^1 (1+x)|g_7(x,t)|dx \\
& + |f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_8(x,t)|dx \Big] \\
& + b_4\Big[2|\hat{w}_{xx}(1,t)| - |\hat{w}_{xx}(0,t)| - \|\hat{w}_{xx}(\cdot,t)\|_{L^1} \\
& + |\dot{\hat{D}}(t)| \int_0^1 (1+x)|g_9(x,t)|dx \\
& + |f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_{10}(x,t)|dx \Big] \\
& + 2\tilde{D}(t)\dot{\hat{D}}(t) \\
& + b_2\dot{\hat{D}}(t) \int_0^1 (1+x)|\hat{w}(x,t)|dx \\
& + b_3\dot{\hat{D}}(t) \int_0^1 (1+x)|\hat{w}_x(x,t)|dx \\
& + b_4\dot{\hat{D}}(t) \int_0^1 (1+x)|\hat{w}_{xx}(x,t)|dx
\end{aligned} \tag{110}$$

We now aim to bound this derivative term. First, we will find it helpful to define the following functional

$$\begin{aligned}
\hat{W}_0(t) =& |X(t)| + \|\tilde{u}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} \\
& + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} \\
& + \|\hat{w}_{xx}(\cdot,t)\|_{L^1}\,.
\end{aligned} \tag{111}$$

Additionally, for the second term in $\dot{\hat{W}}$, there exist $\alpha_{10}, \alpha_{11} \in \mathcal{K}_\infty$ such that

$$\frac{C_2}{2}\left|\frac{\partial f}{\partial u}(X, \alpha_1(X,t))\right| \leq \alpha_{10}(\hat{W}_0(t)) + \alpha_{11}(\epsilon)\,. \tag{112}$$

Furthermore, define $\eta = \min\{\lambda/2, b_0, b_1, b_2, b_3, b_4\}$ for brevity. To bound all of the terms in $\dot{\hat{W}}$ in terms of the system state and control, we will invoke Lemmas 6, 7, 8, 9, 10 along with properties of class $\mathcal{K}_\infty$ functions:

$$\begin{aligned}
\dot{\hat{W}}(t) \leq & -\eta\hat{W}_0(t) \\
& + |\dot{\hat{D}}(t)|\Big[b_0\alpha_{17}(\hat{W}_0) + b_1\big(2\alpha_{24}(\hat{W}_0) + 2\alpha_{24}(\epsilon)\big) \\
& + b_2\big(2\alpha_{19}(\hat{W}_0) + 2\alpha_{19}(\epsilon)\big) \\
& + 2b_3\big(2\alpha_{21}(\hat{W}_0) + 2\alpha_{21}(\epsilon)\big) \\
& + b_3\big(2\alpha_{25}(\hat{W}_0) + 2\alpha_{25}(\epsilon)\big) \\
& + b_4\big(2\alpha_{27}(\hat{W}_0) + 2\alpha_{27}(\epsilon)\big) \\
& + 2(b_2 + b_3 + b_4)\hat{W}_0(t)\Big] \\
& + 2b_4\big(|\dot{\hat{D}}(t)| + |\dot{\hat{D}}(t)|^2 + |\dot{\hat{D}}(t)|^3\big) \\
& + \big(2\alpha_{29}(\hat{W}_0) + 2\alpha_{29}(\epsilon)\big) \\
& + 2b_4|\ddot{D}(t)|\big(2\alpha_{33}(\hat{W}_0) + 2\alpha_{33}(\epsilon)\big) \\
& + |\tilde{D}(t)|\Big[b_0\alpha_{16}(\hat{W}_0) + 2b_1M_2M_3M_1\epsilon + 2b_1\alpha_{18}(\hat{W}_0) \\
& + b_1\big(2\alpha_{23}(\hat{W}_0) + 2\alpha_{23}(\epsilon)\big) \\
& + 2b_4\big(2\alpha_{34}(\hat{W}_0) + 2\alpha_{34}(\epsilon)\big)\Big] \\
& + |\tilde{u}_x(0,t)|\Big[-b_1 + 2b_4\big(2\alpha_{30}(\hat{W}_0) + 2\alpha_{30}(\epsilon)\big)\Big] \\
& + |\hat{w}_x(0,t)|\Big[-b_3 + 2b_4\big(2\alpha_{31}(\hat{W}_0) + 2\alpha_{31}(\epsilon)\big)\Big] \\
& + |\tilde{u}(0,t)|\Big[-b_0 + \alpha_{10}(\hat{W}_0(t)) + \alpha_{11}(\epsilon) \\
& + b_2\big(2\alpha_{20}(\hat{W}_0) + 2\alpha_{20}(\epsilon)\big) \\
& + 2b_3\big(2\alpha_{22}(\hat{W}_0) + 2\alpha_{22}(\epsilon)\big) \\
& + b_3\big(2\alpha_{26}(\hat{W}_0) + 2\alpha_{26}(\epsilon)\big) \\
& + 2b_4\big(2\alpha_{32}(\hat{W}_0) + 2\alpha_{32}(\epsilon)\big) \\
& + b_4\big(2\alpha_{28}(\hat{W}_0) + 2\alpha_{28}(\epsilon)\big)\Big] \\
& + |\hat{w}(0,t)|\Big[-b_2 + \alpha_{10}(\hat{W}_0(t)) + \alpha_{11}(\epsilon)\Big]
\end{aligned}$$

$$+ |\hat{w}_{xx}(0,t)|\, b_4\Big[-1+|\dot{\hat{D}}(t)|\Big]$$
$$+ 2b_2 M_4\epsilon$$
$$+ 2b_1|\tilde{D}(t)|\Big[M_5\Big(|\dot{\hat{D}}(t)|(2\alpha_{21}(\hat{W}_0)+2\alpha_{21}(\epsilon))$$
$$+ |\tilde{u}(0,t)|(2\alpha_{22}(\hat{W}_0)+2\alpha_{22}(\epsilon))\Big)\Big] \quad (113)$$

To handle the cross terms, we use the following two facts:

$$|\tilde{D}(t)| \le 2\overline{D} =: M_6\,, \quad (114)$$
$$|\dot{\hat{D}}(t)| \le \gamma\alpha_{1,D}(\Gamma_0(t)) \le \gamma\alpha_{35}(\hat{W}_0(t))\,, \quad (115)$$

where $\alpha_{35}\in\mathcal{K}_\infty$. Further, the parameters can cancel the boundaries within a local initial region of attraction. Let $R=\hat{W}_0(0)$. Then, given the conditions below hold:

$$b_1 > 2b_4(2\alpha_{30}(R)+2\alpha_{30}(\epsilon))\,, \quad (116)$$
$$b_3 > 2b_4(2\alpha_{31}(R)+2\alpha_{31}(\epsilon))\,, \quad (117)$$
$$b_2 > \alpha_{10}(R)+\alpha_{11}(\epsilon)\,, \quad (118)$$
$$\begin{aligned} b_0 > {}& \alpha_{10}(R)+\alpha_{11}(\epsilon)+b_2(2\alpha_{20}(R)+2\alpha_{20}(\epsilon)) \\ &+2b_3(2\alpha_{22}(R)+2\alpha_{22}(\epsilon))+b_3(2\alpha_{26}(R)+2\alpha_{26}(\epsilon)) \\ &+2b_4(2\alpha_{32}(R)+2\alpha_{32}(\epsilon))+b_4(2\alpha_{28}(R)+2\alpha_{28}(\epsilon)) \\ &2b_1M_5M_6(2\alpha_{22}(R)+2\alpha(\epsilon))\,, \end{aligned} \quad (119)$$

we obtain

$$\begin{aligned} \dot{\hat{W}}(t) \le {}& -\eta\hat{W}_0(t)+|\dot{\hat{D}}(t)|\big[b_0\alpha_{17}(\hat{W}_0) \\ &+b_1(2\alpha_{24}(\hat{W}_0)+2\alpha_{24}(\epsilon)) \\ &+b_2(2\alpha_{19}(\hat{W}_0)+2\alpha_{19}(\epsilon)) \\ &+2b_3(2\alpha_{21}(\hat{W}_0)+2\alpha_{21}(\epsilon)) \\ &+b_3(2\alpha_{25}(\hat{W}_0)+2\alpha_{25}(\epsilon)) \\ &+b_4(2\alpha_{27}(\hat{W}_0)+2\alpha_{27}(\epsilon)) \\ &+2(b_2+b_3+b_4)\hat{W}_0(t)\big] \\ &+2b_4(|\dot{\hat{D}}(t)|+|\dot{\hat{D}}(t)|^2+|\dot{\hat{D}}(t)|^3) \\ &\times(2\alpha_{29}(\hat{W}_0)+2\alpha_{29}(\epsilon)) \\ &+2b_4|\ddot{D}(t)|(2\alpha_{33}(\hat{W}_0)+2\alpha_{33}(\epsilon)) \\ &+|\tilde{D}(t)|\big[b_0\alpha_{16}(\hat{W}_0)+2b_1M_22M_1M_11\epsilon \\ &+2b_1\alpha_{18}(\hat{W}_0) \\ &+b_1(2\alpha_{23}(\hat{W}_0)+2\alpha_{23}(\epsilon)) \\ &+2b_4(2\alpha_{34}(\hat{W}_0)+2\alpha_{34}(\epsilon)) \\ &+2b_1M_5\gamma\alpha_{35}(\hat{W}_0)(2\alpha_{21}(\hat{W}_0)+2\alpha_{21}(\epsilon))\big] \\ &+|\hat{w}_{xx}(0,t)|b_4(-1+|\dot{\hat{D}}(t)|)+2b_2M_2\epsilon \end{aligned} \quad (120)$$

Bounding the perturbation terms Now, we have a series of perturbation terms along with a favorable term $-\eta\hat{W}_0(t)$. Hence, we complete the result by invoking the delay update law, attaining a local result:

$$\begin{aligned} \alpha_{36}(\hat{W}_0(t)) = {}& b_0\alpha_{17}(\hat{W}_0(t))+b_12\alpha_{24}(\hat{W}_0(t)) \\ &+b_22\alpha_{19}(\hat{W}_0(t))+4b_3\alpha_{21}(\hat{W}_0(t)) \\ &+b_32\alpha_{25}(\hat{W}_0(t))+b_42\alpha_{27}(\hat{W}_0(t)) \\ &+2(b_2+b_3+b_4)\hat{W}_0(t)\,, \end{aligned} \quad (121)$$

$$\begin{aligned} \alpha_{37}(\epsilon) = {}& b_12\alpha_{24}(\epsilon) \\ &+b_22\alpha_{19}(\epsilon)+4b_3\alpha_{21}(\epsilon) \\ &+b_32\alpha_{25}(\epsilon)+b_42\alpha_{27}(\epsilon)\,, \end{aligned} \quad (122)$$

$$\begin{aligned} \alpha_{38}(\hat{W}_0(t)) = {}& b_0\alpha_{16}(\hat{W}_0(t)) \\ &+2b_1\alpha_{18}(\hat{W}_0(t))+2b_1\alpha_{23}(\hat{W}_0(t)) \\ &+4b_4\alpha_{34}(\hat{W}_0(t)) \\ &+4b_1M_5\gamma\alpha_{35}(\hat{W}_0(t))\alpha_{21}(\hat{W}_0(t))\,, \end{aligned} \quad (123)$$

$$\alpha_{39}(\epsilon) = 2b_1M_22M_1M_11\epsilon+2b_1\alpha_{23}(\epsilon)+4b_4\alpha_{34}(\epsilon)\,. \quad (124)$$

Using Assumption 4, and the class $\mathcal{K}_\infty$ functions relating Ξ to $\hat{W}_0(t)$, we have $|\dot{\hat{D}}(t)|\le\gamma\alpha_{40}(\hat{W}_0(t))$ and $|\ddot{\hat{D}}(t)|\le\gamma\alpha_{41}(\hat{W}_0(t))$. Let $\gamma\in[0,1]$, then we obtain

$$\begin{aligned} \dot{\hat{W}}(t) \le {}& -\eta\hat{W}_0(t)+\gamma\alpha_{40}(\hat{W}_0(t))(\alpha_{36}(\hat{W}_0(t))+\alpha_{37}(\epsilon)) \\ &+2b_4\gamma(\alpha_{40}(\hat{W}_0(t))+\alpha_{40}^2(\hat{W}_0(t))+\alpha_{40}^3(\hat{W}_0(t))) \\ &\times(2\alpha_{29}(\hat{W}_0(t))+2\alpha_{29}(\epsilon)) \\ &+2b_4\alpha_{41}(\hat{W}_0(t))(2\alpha_{33}(\hat{W}_0(t))+2\alpha_{33}(\epsilon)) \\ &+|\tilde{D}(t)|(\alpha_{38}(\hat{W}_0(t))+\alpha_{39}(\epsilon) \\ &+4b_1M_5\gamma\alpha_{35}(\hat{W}_0(t))\alpha_{21}(\epsilon)) \\ &+|\hat{w}_{xx}(0,t)|b_4\Big[-1+\gamma\alpha_{40}(\hat{W}_0(t))\Big]+2b_2M_2\epsilon\,. \end{aligned} \quad (125)$$

Via Young's inequality, there exists $\alpha_{43:46}\in\mathcal{K}_\infty$ where

$$\begin{aligned} \dot{\hat{W}}(t) \le {}& -\eta\hat{W}_0(t)+\gamma(\alpha_{43}(\hat{W}_0(t))+\alpha_{44}(\epsilon)) \\ &+|\tilde{D}(t)|(\alpha_{45}(\hat{W}_0(t))+\alpha_{46}(\epsilon))+2b_2M_2\epsilon \\ &+|\hat{w}_{xx}(0,t)|b_4\Big[-1+\gamma\alpha_{40}(\hat{W}_0(t))\Big]\,. \end{aligned} \quad (126)$$

Using Young's inequality, for any $\delta>0$ small, we have

$$|\tilde{D}(t)| \le \frac{\delta}{2}+\frac{1}{2\delta}|\tilde{D}^2(t)| \le \frac{\delta}{2}+\frac{1}{2\delta}\hat{W}(t)\,. \quad (127)$$

Hence, we obtain

$$\begin{aligned} \dot{\hat{W}}(t) \le {}& -\eta\hat{W}_0(t)+2b_2M_2\epsilon+\gamma(\alpha_{42}(\hat{W}_0(t))+\alpha_{43}(\epsilon)) \\ &+(\frac{\delta}{2}+\frac{1}{2\delta}\hat{W}(t))(\alpha_{44}(\hat{W}_0(t))+\alpha_{45}(\epsilon)) \\ &+|\hat{w}_{xx}(0,t)|b_4\Big[-1+\gamma\alpha_{40}(\hat{W}_0(t))\Big]\,. \end{aligned} \quad (128)$$

Now, let $\nu\in(0,1)$ and define $C_{R_1} := \max_{x\in[0,R]}\alpha_{42}'(x)$, $C_{R_2} := \max_{x\in[0,R]}\alpha_{44}'(x)$. Then, there exists $\alpha_{46},\alpha_{47}\in\mathcal{K}_\infty$ such that if

$$\gamma<\gamma^* := \min\left\{1, \frac{1}{\max_{x\in[0,R]}\alpha_{40}(x)}, \frac{\eta\nu}{C_{R_1}}\right\}, \quad (129)$$

$$\delta < 2\frac{\eta\nu - \gamma C_{R_1}}{C_{R_2}}, \tag{130}$$

$$\hat{W}(0) < 2\delta\frac{\eta\nu - \gamma C_{R_1} - \delta/2C_{R_2} - \alpha_{46}(\epsilon)}{C_{R_2} + \alpha_{45}(\epsilon)}, \tag{131}$$

then we have

$$\dot{\hat{W}}(t) \le -\eta(1-\nu)\hat{W}_0(t) + \alpha_{47}(\epsilon). \tag{132}$$

Note that $\tilde{D}^2$ is bounded by $(2\overline{\Delta D})^2$. Further, $|X| \le V_0(X) \le \sqrt{C_1}|X|$ by (4). Thus, $\hat{W}_0(t)$ and $\hat{W}$ are equivalent up to $\tilde{D}^2$ such that there exists constants $C_3, C_4 > 0$

$$C_3\hat{W}_0(t) \le \hat{W}(t) \le C_4(\hat{W}_0(t) + \tilde{D}^2(t)). \tag{133}$$

Hence, there exists $\mathcal{K}_\infty$ functions $\alpha_{48}, \alpha_{49}, \alpha_{50}$ such that

$$\dot{\hat{W}}(t) \le -\alpha_{49}(\hat{W}(t)) + \alpha_{50}(\overline{\Delta D}) + \alpha_{51}(\epsilon). \tag{134}$$

From here, using [32, Lemma C.3], we have there exists $\beta_3 \in \mathcal{KL}, \alpha_{52}, \alpha_{53} \in \mathcal{K}_\infty$ such that

$$\hat{W}(t) \le \beta_3(\hat{W}(0), t) + \alpha_{52}(\overline{\Delta D}) + \alpha_{53}(\epsilon). \tag{135}$$

To conclude the result, notice that Υ and $\hat{W}$ are equivalent up to class $\mathcal{K}_\infty$ functions to obtain (89). To obtain the bound on $|X(t)|$, notice that

$$\begin{aligned} |X(t)| &\le W(t) \\ &\le \beta_3(\hat{W}(0), t) + \alpha_{52}(\overline{\Delta D}) + \alpha_{53}(\epsilon), \end{aligned} \tag{136}$$

Using the equivalence between Υ and $\hat{W}$ yields the result with the same exact approach for $\|u(\cdot, t)\|_{L^1}$. ■

7 Numerical Example: Biological Chemostat

To illustrate the theoretical results of Section 6-specifically, neural operator predictors with unknown actuator input-we consider a Chemostat, a bioreactor used for growing micro-organisms or cells under continuous, controlled conditions [49]. Chemostats are widely used in biological engineering, from wastewater treatment [14] to genetic engineering via recombinant DNA [51,37]. Here, we focus on a recent Chemostat model that incorporates population mortality into the dynamics [25]. Consider the nonlinear system

$$\dot{Z}(t) = (\rho_0\mu(S) - \chi - U(t-D))Z(t), \tag{137a}$$

$$\dot{S}(t) = U(t-D)(S_{\text{in}} - S) - \mu(S)Z(t), \tag{137b}$$

where $Z(t) > 0$ is the microbial concentration, $S(t) > 0$ the substrate concentration, $S_{\text{in}} > 0$ the inlet substrate concentration, $U(t) > 0$ the dilution rate, $\rho_0 > 0$ the yield coefficient, μ the population growth rate, $\chi \ge 0$ the mortality rate, and D the dilution delay (i.e. due to valve transport or homogenization time). Without delay, a nominal globally stabilizing controller is given by [25]

$$\begin{aligned} &\kappa(Z(t), S(t)) \\ &\quad = \frac{U^*\mu(S)Z(t)}{\mu(S^*)Z^*} \\ &\quad + \frac{\varsigma \cdot \chi}{(\mu(S^*))^{1+\xi}} \begin{cases} |\mu(S) - \mu(S^*)|^{1+\xi}, & \text{if } S \le S^* \\ 0 & \text{if } S > S^*. \end{cases} \end{aligned} \tag{138}$$

where S^* is the desired substrate concentration, Z^* is the desired microbial population concentration, U^* is the desired dilution rate and $\varsigma > 0$ and $\alpha \in [0, 1)$ are constant feedback gains chosen by the user.

In the experiments that follow, we set the following parameters: $Z^* = 3$, $S^* = 2$, $U^* = 0.9$, $\varsigma = 10$, $\chi = 0.1$, $S_{\text{in}} = 5.33$, $\alpha = 0.5$, $\rho_0 = 1$ and the growth rate function

$$\mu(S) = \frac{7S}{2(1 + S + S^2)}, \tag{139}$$

where we note (Z^*, S^*) is an unstable equilibrium.

We generated a dataset of $100,000$ predictor input–output pairs for training by uniformly varying initial conditions and delay values. Consistent with theory, the unknown actuated input required a smaller ϵ tolerance, and thus more training data and a larger network to achieve stabilization. Among tested architectures, FNO easily stabilizes the system, where as for DeepONet, we deployed the POD-DeepONet variant [40] achieving comparable but slightly weaker practical stability then that of FNO. Figure 2 shows that, without delay compensation, the system oscillates to a limit cycle; yet, with compensation, the learned predictor stabilizes the system. Moreover, Table 2 reports computation costs showing the speedups attained as discretization resolution increases.

Table 2
Computation time (ms) for various approximate predictors averaged over 1000 samples for the Chemostat example.

Step size (dx)	Numerical ↓	DeepONet ↓	FNO ↓	DeepONet Speedup ↑	FNO Speedup ↑
0.01	1.40	**0.34**	1.28	**4.11×**	1.09×
0.005	2.86	**0.43**	1.34	**6.65×**	2.13×
0.001	15.62	**1.02**	1.88	**15.31×**	8.29×

8 Conclusion

We presented the first results on approximate predictors for nonlinear systems with unknown delays, considering cases with measured and unmeasured actuation. In the first setting we established semi-global practical stability, while in the second we obtained local practical stability akin to the exact predictor result. The analysis applies to any uniform predictor, but simulations with neural operators demonstrate stability and significant computational gains over numerical methods in two case

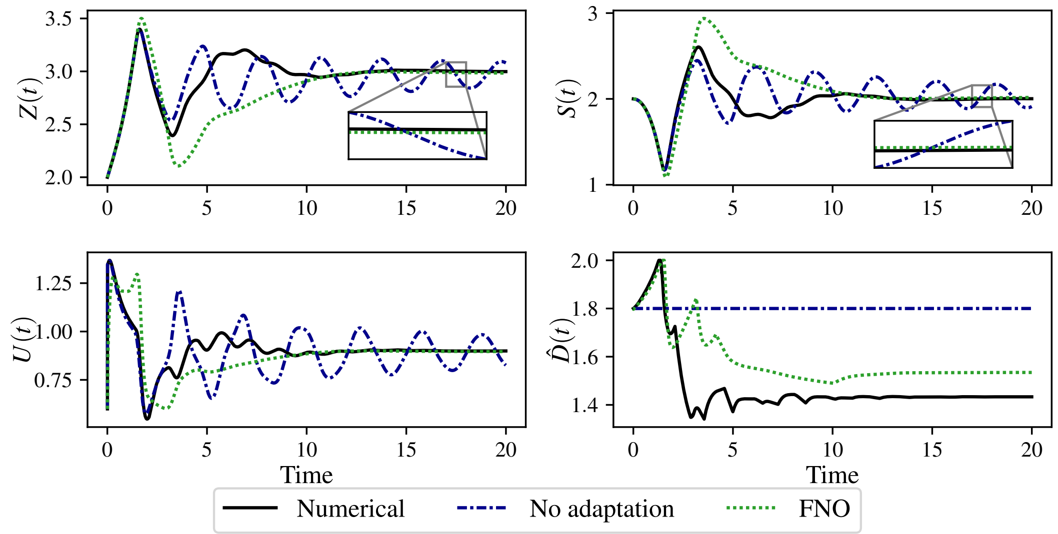

Fig. 2. Chemostat (137) with $Z(0) = 2, S(0) = 2, \hat{D}(0) = 1.8, D = 1.6s$ and target unstable equilibrium $(Z^*, S^*) = (3, 2)$. Comparison of (i) numerical predictor (black), (ii) open-loop without adaptation (blue), and (iii) FNO predictor, stable (green).

studies on an E. coli biological clock and a Chemostat. This work opens directions toward trajectory tracking and output-feedback designs with unknown parameters.

References

[1] Nikolaos Bekiaris-Liberis. Robust string stability and safety of CTH predictor-feedback CACC. *IEEE Transactions on Intelligent Transportation Systems*, 24(8):8209–8221, 2023.

[2] Nikolaos Bekiaris-Liberis and Miroslav Krstic. Compensation of state-dependent input delay for nonlinear systems. *IEEE Transactions on Automatic Control*, 58(2):275–289, 2013.

[3] Nikolaos Bekiaris-Liberis and Miroslav Krstic. *Nonlinear Control Under Nonconstant Delays*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.

[4] Nikolaos Bekiaris-Liberis and Miroslav Krstic. Robustness of nonlinear predictor feedback laws to time- and state-dependent delay perturbations. *Automatica*, 49(6):1576–1590, 2013.

[5] Nikolaos Bekiaris-Liberis and Miroslav Krstic. Predictor-feedback stabilization of multi-input nonlinear systems. *IEEE Transactions on Automatic Control*, 62(2):516–531, 2017.

[6] Alexander Bertino, Peiman Naseradinmousavi, and Miroslav Krstić. Delay-adaptive control of a 7-DOF robot manipulator: Design and experiments. *IEEE Transactions on Control Systems Technology*, 30(6):2506–2521, 2022.

[7] Luke Bhan, Miroslav Krstic, and Yuanyuan Shi. Delay-adaptive control of nonlinear systems with approximate neural operator predictors. In *2025 IEEE 64th Conference on Decision and Control (CDC)*, pages 2085–2090, 2025.

[8] Luke Bhan, Miroslav Krstic, and Yuanyuan Shi. Stabilization of nonlinear systems with unknown delays via delay-adaptive neural operator approximate predictors, 2025. `https://arxiv.org/abs/2509.26443`.

[9] Luke Bhan, Peijia Qin, Miroslav Krstic, and Yuanyuan Shi. Neural operators for predictor feedback control of nonlinear delay systems. In Necmiye Ozay, Laura Balzano, Dimitra Panagou, and Alessandro Abate, editors, *Proceedings of the 7th Annual Learning for Dynamics &; Control Conference*, volume 283 of *Proceedings of Machine Learning Research*, pages 179–193. PMLR, 04–06 Jun 2025.

[10] Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Neural operators for bypassing gain and control computations in PDE backstepping. *IEEE Transactions on Automatic Control*, 69(8):5310–5325, 2024.

[11] Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Adaptive control of reaction–diffusion PDEs via neural operator-approximated gain kernels. *Systems & Control Letters*, 195:105968, 2025.

[12] Delphine Bresch-Pietri and Miroslav Krstic. Delay-Adaptive Control for Nonlinear Systems. *IEEE Transactions on Automatic Control*, 59(5):1203–1218, 2014.

[13] Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks*, 6(4):911–917, 1995.

[14] G. D'Ans, P. Kokotovic, and D. Gottlieb. A nonlinear regulator problem for a model of biological waste treatment. *IEEE Transactions on Automatic Control*, 16(4):341–347, 1971.

[15] D. Del Vecchio. Design and analysis of an activator-repressor clock in e. coli. In *2007 American Control Conference*, pages 1589–1594, 2007.

[16] Florian Dörfler and Francesco Bullo. Synchronization and transient stability in power networks and non-uniform kuramoto oscillators. In *Proceedings of the 2010 American Control Conference*, pages 930–937, 2010.

[17] Nicolás Espitia and Wilfrid Perruquetti. Predictor-feedback prescribed-time stabilization of LTI systems with input delay. *IEEE Transactions on Automatic Control*, 67(6):2784–2799, 2022.

[18] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*, volume 18 of *Mathematics and Its Applications (Soviet Series)*. Kluwer Academic Publishers, Dordrecht, 1988.

[19] Michael A Henson and Dale E Seborg. Time delay compensation for nonlinear processes. *Industrial & engineering chemistry research*, 33(6):1493–1500, 1994.

[20] Fayçal Ikhouane and Miroslav Krstić. Adaptive backstepping with parameter projection: Robustness and asymptotic performance. *Automatica*, 34(4):429–435, 1998.

[21] M. Jankovic. Control Lyapunov-Razumikhin functions and robust stabilization of time delay systems. *IEEE Transactions on Automatic Control*, 46(7):1048–1060, 2001.

[22] M. Jankovic. Control of nonlinear systems with time delay. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, volume 5, pages 4545–4550 Vol.5, 2003.

[23] Iasson Karafyllis. Finite-time global stabilization by means of time-varying distributed delay feedback. *SIAM Journal on Control and Optimization*, 45(1):320–342, 2006.

[24] Iasson Karafyllis and Miroslav Krstic. *Predictor Feedback for Delay Systems: Implementations and Approximations*. Birkhäuser Cham, 2017.

[25] Iasson Karafyllis, Epiphane Loko, Miroslav Krstic, and Antoine Chaillet. Global stabilization of chemostats with nonzero mortality and substrate dynamics, 2025.

[26] Hassan K Khalil. *Nonlinear systems*. Prentice Hall, Upper Saddle River, N.J., 2002.

[27] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.

[28] Nikola B. Kovachki, Samuel Lanthaler, and Hrushikesh Mhaskar. Data Complexity Estimates for Operator Learning, 2024.

[29] Costas Kravaris and Raymond A. Wright. Deadtime compensation for nonlinear processes. *AIChE Journal*, 35(9):1535–1542, 1989.

[30] Miroslav Krstic. Input Delay Compensation for Forward Complete and Strict-Feedforward Nonlinear Systems. *IEEE Transactions on Automatic Control*, 55(2):287–303, 2010.

[31] Miroslav Krstic, Luke Bhan, and Yuanyuan Shi. Neural operators of backstepping controller and observer gain functions for reaction–diffusion PDEs. *Automatica*, 164:111649, 2024.

[32] Miroslav Krstic, Petar V. Kokotovic, and Ioannis Kanellakopoulos. *Nonlinear and Adaptive Control Design*. John Wiley & Sons, Inc., USA, 1st edition, 1995.

[33] Maxence Lamarque, Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Adaptive neural-operator backstepping control of a benchmark hyperbolic pde. *Automatica*, 177:112329, 2025.

[34] Maxence Lamarque, Luke Bhan, Rafael Vazquez, and Miroslav Krstic. Gain scheduling with a neural operator for a transport PDE with nonlinear recirculation. *IEEE Transactions on Automatic Control*, pages 1–8, 2025.

[35] Samuel Lanthaler, Zongyi Li, and Andrew M. Stuart. Nonlocality and nonlinearity implies universality in operator learning. *Constructive Approximation*, 62(2):261–303, Oct 2025.

[36] Samuel Lanthaler, Andrew M. Stuart, and Margaret Trautner. Discretization Error of Fourier Neural Operators, 2024.

[37] B R Levin and F M Stewart. The population biology of bacterial plasmids: a priori conditions for the existence of mobilizable nonconjugative factors. *Genetics*, 94(2):425–443, February 1980.

[38] Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021.

[39] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, Mar 2021.

[40] Lu Lu, Xuhui Meng, Shengze Cai, Zhiping Mao, Somdatta Goswami, Zhongqiang Zhang, and George Em Karniadakis. A comprehensive and fair comparison of two neural operators (with practical extensions) based on fair data. *Computer Methods in Applied Mechanics and Engineering*, 393:114778, 2022.

[41] Yuanyuan shi Luke Bhan, Miroslav Krstic. Stabilization of nonlinear systems with unknown delays via delay-adaptive neural operator approximate predictors. GitHub repository, 2026. Available at: `https://github.com/lukebhan/Neural-Operator-Delay-Adaptive-Predictor-Feedback`, accessed April 8, 2026.

[42] F. Mazenc and P.-A. Bliman. Backstepping design for time-delay nonlinear systems. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, volume 5, pages 4551–4556 Vol.5, 2003.

[43] Tamas G. Molnar, Adam K. Kiss, Aaron D. Ames, and Gábor Orosz. Safety-critical control with input delay in dynamic environment. *IEEE Transactions on Control Systems Technology*, 31(4):1507–1520, 2023.

[44] Anirbit Mukherjee and Amartya Roy. Size Lowerbounds for Deep Operator Networks. *Transactions on Machine Learning Research*, 2024.

[45] Marios M. Polycarpou and Petros A. Ioannou. A robust adaptive nonlinear control design. In *1993 American Control Conference*, pages 1365–1369, 1993.

[46] M.M. Polycarpou. Stable adaptive neural control scheme for nonlinear systems. *IEEE Transactions on Automatic Control*, 41(3):447–451, 1996.

[47] Joseph R. Pomerening, Eduardo D. Sontag, and James E. Ferrell. Building a cell cycle oscillator: hysteresis and bistability in the activation of cdc2. *Nature Cell Biology*, 5(4):346–351, Apr 2003.

[48] Jie Qi, Jing Zhang, and Miroslav Krstic. Neural operators for PDE backstepping control of first-order hyperbolic PIDE with recycle and delay. *Systems & Control Letters*, 185:105714, 2024.

[49] Hal L. Smith and Paul Waltman. *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge Studies in Mathematical Biology. Cambridge University Press, 1995.

[50] Otto JM Smith. Closer control of loops with dead time. *Chemical engineering progress*, 53:217–219, 1957.

[51] Gregory Stephanopoulos and Gary R. Lapudis. Chemostat dynamics of plasmid-bearing, plasmid-free mixed recombinant cultures. *Chemical Engineering Science*, 43(1):49–57, 1988.

[52] Samir Toumi. Neural operators in backstepping control for hyperbolic partial differential equations (PDEs) in rotary drilling systems. *Systems & Control Letters*, 204:106191, 2025.

[53] Shanshan Wang, Mamadou Diagne, and Miroslav Krstić. Deep learning of delay-compensated backstepping for reaction-diffusion PDEs. *IEEE Transactions on Automatic Control*, pages 1–8, 2025.

[54] Yu Xiao, Yuan Yuan, Biao Luo, and Xiaodong Xu. Neural operators for robust output regulation of hyperbolic PDEs. *Neural Networks*, 179:106620, 2024.

[55] Xiang Xu, Lu Liu, and Gang Feng. Stability and stabilization of infinite delay systems: A Lyapunov-based approach. *IEEE Transactions on Automatic Control*, 65(11):4509–4524, 2020.

[56] Yihuai Zhang, Ruiguo Zhong, and Huan Yu. Neural operators for boundary stabilization of stop-and-go traffic. In Alessandro Abate, Mark Cannon, Kostas Margellos, and Antonis Papachristodoulou, editors, *Proceedings of the 6th Annual Learning for Dynamics &; Control Conference*, volume 242 of *Proceedings of Machine Learning Research*, pages 554–565. PMLR, 15–17 Jul 2024.

A Lemma's used in Section 4

Proofs for Lemmas 4-11 are available in [8].

Lemma 4 *There exist class $\mathcal{K}_\infty$ functions α_{12}, α_{13} such that, for all $x \in [0,1]$,*

$$|\hat{p}(x,t)| \le \alpha_{12}\big(|X| + \|\hat{u}(\cdot,t)\|_{L^1}\big)\,, \tag{A.1}$$
$$|\hat{p}(x,t)| \le \alpha_{13}\big(|X| + \|\hat{w}(\cdot,t)\|_{L^1}\big)\,. \tag{A.2}$$

Lemma 5 *There exists $\mathcal{K}_\infty$ function α_{14}, α_{15} such that*

$$|\hat{u}(x,t)| \le \alpha_{14}(|X| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1}) + M_1\epsilon\,, \tag{A.3}$$
$$|\tilde{u}(x,t)| \le \alpha_{15}(\|\tilde{u}_x(\cdot,t)\|_{L^1})\,. \tag{A.4}$$

Lemma 6 *There exists $\alpha_{16}, \alpha_{17}, \alpha_{18} \in \mathcal{K}_\infty$ and constant $M_5 = \frac{1}{\underline{D}}$ such that*

$$\int_0^1 (1+x)|g_1(x,t)|dx \le \alpha_{16}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1})\,, \tag{A.5}$$
$$\int_0^1 (1+x)|g_2(x,t)|dx \le \alpha_{17}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1})\,, \tag{A.6}$$
$$|\tilde{u}_x(1,t)| \le |\tilde{D}(t)|\big[M_5|\hat{w}_x(1,t)| + M_2M_1M_11\epsilon + \alpha_{18}(|X| + \|\hat{w}(\cdot,t)\|_{L^1})\big]\,. \tag{A.7}$$

Lemma 7 *There exists $\alpha_{19:22} \in \mathcal{K}_\infty$ such that*

$$\int_0^1 (1+x)|g_3(x,t)|dx \le \alpha_{19}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon)\,, \tag{A.8}$$
$$|f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_4(x,t)|dx \le |\tilde{u}(0,t)| \times \alpha_{20}(|X(t)| + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon)\,, \tag{A.9}$$
$$|\hat{w}_x(1,t)| \le |\dot{\hat{D}}(t)|\alpha_{21}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon) + \tilde{u}(0,t)| \times \alpha_{22}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon)\,. \tag{A.10}$$

Lemma 8 *There exists α_{23}, α_{24} in $\mathcal{K}_\infty$ such that*

$$\int_0^1 (1+x)|g_5(x,t)|dx \le \alpha_{23}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \epsilon)\,, \tag{A.11}$$
$$\int_0^1 (1+x)|g_6(x,t)|dx \le \alpha_{24}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \epsilon)\,. \tag{A.12}$$

Lemma 9 *There exists $\alpha_{25}, \alpha_{26} \in \mathcal{K}_\infty$ such that*

$$\int_0^1 (1+x)|g_7(x,t)|dx \le \alpha_{25}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \epsilon)\,, \tag{A.13}$$
$$|f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_8(x,t)|dx \le |\tilde{u}(0,t)| \times \alpha_{26}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\tilde{u}_x[t]\|_{L^1} + \epsilon)\,. \tag{A.14}$$

Lemma 10 *There exists $\alpha_{27}, \alpha_{28} \in \mathcal{K}_\infty$ such that*

$$\int_0^1 (1+x)|g_9(x,t)|dx \le |\hat{w}_{xx}(0,t)| + \alpha_{27}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \epsilon)\,, \tag{A.15}$$
$$|f_{\tilde{u}}(t)| \int_0^1 (1+x)|g_{10}(x,t)|dx \le |\tilde{u}(0,t)| \times \alpha_{28}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1})\,. \tag{A.16}$$

Lemma 11 *There exists $\alpha_{29:34} \in \mathcal{K}_\infty$ such that*

$$\begin{aligned}
|\hat{w}_{xx}(1,t)| \le &\Big(|\dot{\hat{D}}(t)| + |\dot{\hat{D}}(t)^2| + |\dot{\hat{D}}(t)^3|\Big) \times \alpha_{29}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \|\hat{w}_{xx}(\cdot,t)\|_{L^1} + \epsilon) \\
&+ |\tilde{u}_x(0,t)|\alpha_{30}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon) \\
&+ |\hat{w}_x(0,t)|\alpha_{31}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon) \\
&+ |\tilde{u}(0,t)|\alpha_{32}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon) \\
&+ |\ddot{D}(t)|\alpha_{33}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon) \\
&+ |\tilde{D}(t)|\alpha_{34}(|X(t)| + \|\hat{w}(\cdot,t)\|_{L^1} + \|\tilde{u}_x(\cdot,t)\|_{L^1} + \|\hat{w}_x(\cdot,t)\|_{L^1} + \epsilon)\,.
\end{aligned} \tag{A.17}$$