Automatica 177 (2025) 112329

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Adaptive neural-operator backstepping control of a benchmark hyperbolic PDE[☆]

Maxence Lamarque[a], Luke Bhan[b,*], Yuanyuan Shi[b], Miroslav Krstic[b]

[a] École des Mines de Paris, France
[b] University of California, San Diego, USA

ARTICLE INFO

Article history:
Received 15 January 2024
Received in revised form 28 January 2025
Accepted 19 March 2025

Keywords:
Partial differential equations
Adaptive control
Deep learning
Lyapunov stability

ABSTRACT

In this paper, we develop the first result employing neural operators in adaptive PDE control, presented for a benchmark 1-D hyperbolic PDE with recirculation. Particularly, we introduce neural operators for approximating the mapping from the adaptive estimation of the plants' functional coefficients to the corresponding controller gain kernel. This nonlinear mapping is computationally prohibitive in adaptive control when the resulting gain kernel needs to be continuously resolved as the estimation of the plant functional coefficient is updated. Thus, by introducing a neural operator approximation of this mapping, we absolve the computational barrier for implementing real-time adaptive control of PDEs. We establish global stabilization via Lyapunov analysis, in the plant and parameter error states, and also present an alternative approach, via passive identifiers, which avoids the strong assumptions on kernel differentiability. We then present numerical simulations demonstrating stability and observe speedups up to three orders of magnitude, highlighting the real-time efficacy of neural operators in adaptive control. Our code (Github) is made publicly available for future researchers.

1. Introduction

Following several papers in which PDE backstepping controllers were shown *robust* to the implementation of the gain kernels by replacing the solution of kernel PDEs by an offline-computed neural operator (NO) approximation of the kernel (Bhan, Shi, & Krstic, 2023; Krstic, Bhan, & Shi, 2023; Qi, Zhang, & Krstic, 2023; Wang, Diagne, & Krstić, 2023a, 2023b; Zhang, Zhong, & Yu, 2023), in this paper we introduce the *first adaptive* backstepping controller where the gain kernels are computed via NOs in real time, from online parameter estimates. We do so for a hyperbolic PDE with linear recirculation, the most accessible but nevertheless nontrivial (unstable) PDE system, with a functional coefficient that is unknown, and with boundary actuation.

We employ an (indirect) adaptive version of a standard PDE backstepping controller for a 1-D hyperbolic PDEs but with the analytical gain kernel replaced with the operator approximated equivalent. We then show, under the kernel operator approximation, global stability of the resulting closed-loop system via Lyapunov analysis and neural operator approximation theorems (Deng, Shin, Lu, Zhang, & Karniadakis, 2022; Lanthaler, Li, & Stuart, 2023). Furthermore, we present an alternative approach based on passive identifiers simplifying the assumptions on the gain-kernel derivatives at a cost of an increased dynamic order of the parameter estimator.

This is the first result in which *offline learning* and *online learning* are both employed, working in tandem. Hence, it is important to explain these two distinct learning tasks. The operator from the plant coefficient to the kernel is learned offline — once and for all. The unknown plant coefficient is learned online, continually, using a parameter estimator. The offline and online learners are combined through the adaptive gain, where the NO is evaluated, at each time step, for the new plant coefficient estimate. The NO speeds up the evaluation of the adaptive gain by about $10^3\times$, relative to the hypothetical online solving of the gain kernel equation, and thus enables the real-time adaptive control of the PDE.

Given the value of the $10^3\times$ speedup in computing the adaptive gain, the code for all the computational tasks performed in relation to this adaptive design are made publicly available at https://github.com/lukebhan/NeuralOperatorAdaptiveControl.

Stabilization of PDEs using backstepping-based adaptive control. The first investigations into backstepping-based adaptive control of PDEs were introduced for reaction–diffusion PDEs. Initially, a set of three approaches extending the simpler ODE counterparts were introduced: a Lyapunov approach (Krstic &

☆ The material in this paper was presented at None. This paper was recommended for publication in revised form by Associate Editor Thomas Meurer under the direction of Editor Luca Zaccarian.

* Corresponding author.
E-mail addresses: maxence.lamarque@etu.minesparis.psl.eu (M. Lamarque), lbhan@ucsd.edu (L. Bhan), yyshi@ucsd.edu (Y. Shi), krstic@ucsd.edu (M. Krstic).

https://doi.org/10.1016/j.automatica.2025.112329
0005-1098/

Smyshlyaev, 2008), a passive identification approach where one constructs an observer-like PDE system to estimate the plant parameter (Smyshlyaev & Krstic, 2007a), and a swapping identifier where filters are introduced for the measurement to create a prediction error which can be minimized via standard techniques such as gradient descent (Smyshlyaev & Krstic, 2007b). Papers Bresch-Pietri and Krstic (2009), Bresch-Pietri and Krstic (2014) then extended these techniques to adaptive control for systems with unknown delays in ODEs, and to wave PDEs. This paved the way for a swapping-based output-feedback extension to a single hyperbolic PDE (Bernard & Krstic, 2014), and then to extensions to systems of hyperbolic PDES (Anfinsen & Aamo, 2018, 2019; Anfinsen, Diagne, Aamo, & Krstic, 2016). Further, Zhu and Krstic (2020) expanded the direction introduced by Bresch-Pietri and Krstic (2009) into a series of works on adaptive control of delay-systems. Concurrently, many works explored adaptive backstepping for different systems including coupled hyperbolic PDEs in Yu, Vazquez, and Krstic (2017), coupled hyperbolic PDE-PDE-ODE systems in Wang, Tang, and Krstic (2023), and the wave equation in Wang, Tang, and Krstic (2020). Lastly, we briefly mention the more recent works in adaptive control expanding into distributed input systems with unknown delays (Wang, Diagne, & Qi, 2022) and event-triggered adaptive control of coupled hyperbolic PDEs (Karafyllis, Krstic, & Chrysafi, 2019; Wang & Krstic, 2021, 2023).

Neural operator approximations for model-based PDE control. In a series of breakthrough innovations in the mathematics of machine learning (Lanthaler et al., 2023; Lanthaler, Mishra, & Karniadakis, 2022; Lu, Jin, Pang, Zhang, & Karniadakis, 2021), universal operator approximation theorems have been developed which demonstrate that neural networks can effectively approximate mappings across function spaces. Naturally, the control community then capitalized on these results to approximate the kernel operator in PDE Backstepping. The first study in this direction was conducted for a 1D transport PDE in Bhan et al. (2023), and then later extended to both a reaction–diffusion PDE and observers in Krstic et al. (2023). In both works, the stability of the PDE under the approximated kernel is rigorously proved by employing the universal operator approximation theorem (Deng et al., 2022). Following Bhan et al. (2023), Krstic et al. (2023), there have been a series of extensions where Qi et al. (2023), Wang et al. (2023a) developed similar results for hyperbolic and parabolic PDEs with delays. Paper Zhang et al. (2023) then tackles the first application of NO approximations controlling the Aw–Rascale–Zhang(ARZ) PDE consisting of a set of second-order coupled hyperbolic PDEs describing traffic flows. Furthermore, Wang et al. (2023b) then considers NOs for a more general form of 2 × 2 hyperbolic PDEs with applications to oil drilling and shallow water wave modeling. Lastly, Lamarque, Bhan, Vazquez, and Krstic (2024) employ neural operators for gain-scheduling of hyperbolic PDEs with nonlinear recirculation — the first of such work where the kernel is recomputed at every timestep thus enabling real-time control of nonlinear PDEs.

Contributions. Two major advances in methodology and analysis are made. For Lyapunov-based and observer-based (passive) designs of update laws, two distinct neural operators are employed. For the Lyapunov update, a smoother NO is trained (the so-called "full-kernel" NO), leading to a target system with a homogeneous boundary condition and perturbations in the domain, whereas for the observer-based update, introduced in Anfinsen and Aamo (2019), a simpler but less smooth NO is trained (the so-called "gain-only" NO), eliminating the perturbation in the PDE's domain but making the boundary condition perturbed. These two designs give rise to distinct mathematical issues to overcome. The paper not only solves the technical problems that arise in NO-based adaptive PDE control but also illuminates the tradeoff between the two NO approaches.

Table 1
Nomenclature for offline and online kernel learning.

Exact operator	$\mathscr{K}$
Neural (approximate) operator	$\hat{\mathscr{K}}$
Exact kernel	$k = \mathscr{K}(\beta)$
Exact estimated kernel	$\breve{k} = \mathscr{K}(\hat{\beta})$
Approximate estimated kernel (adaptive kernel)	$\hat{k} = \hat{\mathscr{K}}(\hat{\beta})$

The key novel mathematical challenge overcome in this paper, relative to the papers Bhan et al. (2023), Krstic et al. (2023), Qi et al. (2023), Wang et al. (2023a, 2023b), Zhang et al. (2023) in which the robustness to NO approximating of the gain is established, is that the updating of the plant coefficient, and the associated updating of the kernel through the NO, gives rise to not only a potentially high rate of change in the adaptive gain but also a potentially high rate of change of the error in the NO approximation of the adaptive gain. This mathematical challenge is handled differently in the Lyapunov/full-kernel and observer-based/gain-only approaches. Each approach has its merit and each of the proof procedures has an educational value to the reader aspiring to pursue extensions of NO-enabled adaptive control of PDEs.

The most obvious contribution is in the enablement of real-time adaptive PDE control, through a $10^3\times$ speedup in the computation of the adaptive gain.

Paper outline. In Section 2, we briefly restate the unpublished but relatively easy result for adaptive PDE backstepping of hyperbolic PDEs with recirculation. In Section 3, we prove both existence and boundedness of the exact backstepping kernel and its derivative. In Section 4, we then present the neural operator approximation theorem and show the adaptive backstepping kernel can be approximated by a neural operator. Next, in Section 5, we give the paper's main result presenting stability of the closed loop feedback system under the neural operator. We follow the result with a proof in Section 6 via Lyapunov analysis. In Section 7, we present an alternative approach, via a modular design with a passive identifier that avoids the approximation of the kernel's derivative and thus the strong assumptions about the kernel's differentiability required for Lyapunov analysis. Lastly, in Section 8, we present numerical simulations highlighting the theoretical stability result and calculate the numerical speedups gained from the neural operator approximation.

Notation. We present the nomenclature for the offline and online kernel in Table 1. We abbreviate the partial derivative as $a_x(x,\nu) = \frac{\partial a}{\partial x}(x,\nu)$. For a function a defined on $[0,1]\times\mathbb{R}^+$ we denote the spatial L^2 norm as $\|a(t)\| = \sqrt{\int_0^1 a^2(x,t)dx}$, which is a function of $t\in\mathbb{R}^+$. We write $f(c) = \mathscr{O}_{c\to\infty}(g(c))$ if there exists $c_0 \ge 0$ such that for all $c \ge c_0$, $|f(c)| \le M|g(c)|$ for some uniform constant $M > 0$. We denote the convolution operation (which is commutative) by

$$(a*b)(x,t) = (b*a)(x,t) = \int_0^x a(x-y,t)b(y,t)dy\,. \tag{1}$$

2. Exact adaptative PDE backstepping for a hyperbolic PDE with recirculation

We consider the following hyperbolic PDE—transport PDE with recirculation,

$$u_t(x,t) = u_x(x,t) + \beta(x)u(0,t)\,, \quad \forall (x,t)\in[0,1)\times\mathbb{R}^+ \tag{2}$$

$$u(1,t) = U(t)\,, \tag{3}$$

where β is an unknown function to be estimated online using an estimate $\hat{\beta}(x,t)$.

We employ an *adaptive backstepping transformation* given by

$$w(x,t) = u(x,t) - \breve{k} * u(x,t)\,, \qquad \forall (x,t) \in [0,1]\times\mathbb{R}\,, \tag{4}$$

where $\breve{k}$ is the (online) backstepping kernel solution of the Volterra equation

$$\breve{k}(x,t) = -\hat{\beta}(x,t) + \hat{\beta} * \breve{k}(x,t)\,, \quad (x,t) \in [0,1]\times\mathbb{R}^+\,. \tag{5}$$

The transformation (4) maps the system (2), (3) to the perturbed target system

$$\begin{aligned} w_t(x,t) &= w_x(x,t) \\ &\quad + \left[\tilde{\beta}(x,t) - \breve{k} * \tilde{\beta}(x,t)\right] w(0,t) - \Omega(x,t)\,, \end{aligned} \tag{6}$$

$$w(1,t) = 0, \tag{7}$$

where

$$\tilde{\beta}(x,t) = \beta(x,t) - \hat{\beta}(x,t), \tag{8}$$

$$\Omega(x,t) = \breve{k}_t * (w - \breve{l} * w)(x,t)\,, \tag{9}$$

$$\breve{l} = -\breve{k} + \breve{k} * \breve{l} = \mathscr{K}\left(\breve{k}\right) = \mathscr{K}\circ\mathscr{K}\left(\hat{\beta}\right) = \hat{\beta}\,. \tag{10}$$

Note that the boundary condition (6) gives from (4) the feedback law

$$U(t) = (\breve{k} * u)(1,t) = (\mathscr{K}(\hat{\beta}) * u)(1,t). \tag{11}$$

We first state an adaptive control design for the adaptative problem with the exact backstepping kernel $\breve{k}$. The next theorem serves only as a guidance for what we seek to achieve under a *NO-based approximate adaptive backstepping* design. We omit the theorem's proof since it can be deduced from the proof of our main result in Theorem 4.

Theorem 1 (*Full-State exact Adaptative Control Design*). *Consider the plant (2)–(3) in feedback with the control law*

$$U(t) = \int_0^1 \breve{k}(1-y,t)u(y,t)dy, \quad t \geq 0 \tag{12}$$

where $\breve{k}$ is solution of the Volterra integral equation (5). For all $c > 0$ and all $B > 0$ such that $\|\beta\|_\infty \leq B$, there exists $\gamma^*(c,B) = \mathscr{O}_{c\to\infty}(e^{-c}) > 0$ with a decreasing dependence on B, such that for all $\gamma \in (0,\gamma^*)$, any initial condition $\hat{\beta}(\cdot,0) \in \mathscr{C}^0([0,1],\mathbb{R})$ satisfying $\|\hat{\beta}(\cdot,0)\|_\infty \leq B$, the update law

$$\hat{\beta}_t(x,t) := \mathrm{Proj}(\tau(x,t),\hat{\beta}(x,t)), \quad \forall (x,t)\in[0,1]\times\mathbb{R}^+\,, \tag{13}$$

$$\tau(x,t) := \frac{\gamma}{1+\|w(t)\|_c^2}\left[e^{cx}w(x,t) - \int_x^1 \breve{k}(y-x,t)e^{cy}w(y,t)dy\right] w(0,t), \tag{14}$$

where

$$w = u - \breve{k} * u\,, \tag{15}$$

$$\|w(t)\|_c^2 = \int_0^1 e^{cx}w^2(x,t)dx\,, \tag{16}$$

with the projection operator $\mathrm{Proj} : \mathbb{R}\times[0,B]\to\mathbb{R}$ *defined as*[1]

$$\mathrm{Proj}(a,b) := \begin{cases} 0, & \text{if } |b| = B \text{ and } ab > 0 \\ a, & \text{else} \end{cases} \tag{17}$$

guarantees that

$$\Gamma(t) \leq R(e^{\rho\Gamma(0)} - 1)\,, \qquad \forall t \geq 0, \tag{18}$$

$$\Gamma(t) = \int_0^1 \left[u^2(x,t) + \left(\beta(x) - \hat{\beta}(x,t)\right)^2\right] dx \tag{19}$$

for constants $\rho, K > 0$ and, in addition, $u(x,t) \underset{t\to\infty}{\to} 0$ for all $x \in [0,1]$.

In summary, with the *exact* adaptive backstepping feedback law (12), (5), (13), (14), the equilibrium $(u(x),\hat{\beta}(x)) \equiv (0,\beta(x))$ is globally stable in the L^2 sense and the state $u(x,t)$ is regulated to zero pointwise in x. The computationally intensive part of implementing this feedback law is that the Volterra equation (5) needs to be solved (in x) at each time "step" t. It is for this reason that we seek a neural operator approximation $\hat{\mathscr{K}} : \hat{\beta} \mapsto \hat{k}$ to the exact adaptive backstepping gain operator $\mathscr{K} : \hat{\beta} \mapsto \breve{k}$, which would require only a neural network *evaluation* at each t, rather than a solution to a Volterra equation.

3. Backstepping kernel properties

This section introduces results on the exact adaptive backstepping kernel $\breve{k}$ in (5).

Lemma 1 (*Existence and Upper Bound for Kernel and Its Derivative*). *Let $B > 0$, $\hat{\beta} \in \mathscr{C}^0([0,1]\times\mathbb{R}^+,\mathbb{R})$ such that $\|\hat{\beta}\|_\infty \leq B$ and consider the Volterra equation (5), reiterated here for convenience,*

$$\breve{k}(x,t) - \hat{\beta} * \breve{k}(x,t) + \hat{\beta}(x,t) = 0\,, \quad (x,t)\in[0,1]\times\mathbb{R}^+. \tag{20}$$

There exists a unique $\mathscr{C}^0([0,1]\times\mathbb{R}^+,\mathbb{R})$ solution $\breve{k}$ that satisfies

$$\|\breve{k}\|_\infty \leq Be^B. \tag{21}$$

If, in addition, $\hat{\beta}_t$ exists and is continuous with respect to x on $[0,1]\times\mathbb{R}^+$ such that

$$\|\hat{\beta}_t\|_{\infty,[0,1]\times[0,T]} < \infty, \quad \forall T > 0, \tag{22}$$

then $\breve{k}_t$ exists, is continuous with respect to x on $[0,1]\times\mathbb{R}^+$, and satisfies

$$\|\breve{k}_t(t)\| \leq \|\hat{\beta}_t(t)\|(1 + Be^B(2+Be^B)), \quad t \geq 0\,. \tag{23}$$

Proof. Let $B > 0$, $\hat{\beta} \in \mathscr{C}^0([0,1]\times\mathbb{R}^+,\mathbb{R})$ such that $\|\hat{\beta}\|_\infty \leq B$. We notice that (20) is just a Volterra integral equation since $\hat{\beta}$ is continuous. The existence and continuity of $\breve{k}$ follows. Also, note that (20) implies

$$|\breve{k}(x,t)| \leq B + \int_0^x |\hat{\beta}(x-y,t)|.|\breve{k}(y,t)|dy. \tag{24}$$

Then Grönwall's lemma gives (21). We now prove the existence and continuity with respect to x of $\breve{k}_t$ on $[0,1]\times\mathbb{R}^+$. To do so we use a successive approximation approach. We introduce the sequence

$$\Delta k^0 := -\hat{\beta}\,, \tag{25}$$

$$\Delta k^{n+1} := \hat{\beta} * \Delta k^n\,. \tag{26}$$

Through iteration we have

$$|\Delta k^n(x,t)| \leq \frac{B^{n+1}x^n}{n!}, \quad (x,t)\in[0,1]\times\mathbb{R}^+\,, \tag{27}$$

From which we have

$$\breve{k} = \sum_{n=0}^{\infty} \Delta k^n\,. \tag{28}$$

[1] The projector operator defined here is not continuous. Hence, the solutions of the PDE system are in the Filippov sense. To avoid the discontinuity, one would add a boundary layer of width $\delta > 0$. But to avoid having the exposition drifting into inessential technicalities, we use the common discontinuous projection (17).

We will be proving that the series $\sum_{n=0}^{\infty} \Delta k_t^n$ uniformly converges on each compact $[0,1]\times[0,T]$, $T>0$. We begin by introducing the function

$$B_t(T) := \|\hat{\beta}_t\|_{\infty,[0,1]\times[0,T]} < \infty, \quad T>0, \tag{29}$$

with assumption (29). We can prove through induction that $\forall n\in\mathbb{N}$, Δk_t^n exists, is continuous with respect to x, and satisfies

$$|\Delta k_t^n(x,t)| \le \frac{(n+1)\alpha(T)x^n}{n!}, \quad \forall (x,t)\in[0,1]\times[0,T], \tag{30}$$

$$\alpha(T) := \max\{B_t(T), B\}. \tag{31}$$

To do so we have to take the derivative of (26) with respect to t using the Leibniz theorem for the derivation of integral with parameter. Notice that we use the 'strong' version of this theorem that does not require $\hat{\beta}_t$ to be continuous with respect to t. From (30), we have that the series $\sum_{n=0}^{\infty}\Delta k_t^n$ uniformly converges on $[0,1]\times[0,T]$, for all $T>0$. We thus have the existence and continuity with respect to x of $\breve{k}_t$ on $[0,1]\times\mathbb{R}^+$. We can then take the derivative of (20) with respect to t, which gives the following inequality satisfied by $\breve{k}_t$,

$$\begin{aligned}|\breve{k}_t(x,t)| \le{}& |\hat{\beta}_t(x,t)| + Be^B\|\hat{\beta}_t(x,t)\| \\ &+ \int_0^x |\hat{\beta}(x-y,t)||\breve{k}_t(y,t)|dy, \\ &\forall (x,t)\in[0,1]\times\mathbb{R}^+.\end{aligned} \tag{32}$$

Then using Grönwall's lemma on (32), we arrive at

$$\begin{aligned}|\breve{k}_t(t,x)| \le{}& |\hat{\beta}_t(x,t)| + Be^B\|\hat{\beta}_t(t)\|(2+Be^B), \\ &\forall (x,t)\in[0,1]\times\mathbb{R}^+.\end{aligned} \tag{33}$$

From (33), using the triangular inequality we have (23). □

4. Neural operator approximation of backstepping kernel

Explicitly solving the Volterra equation (5) satisfied by $\breve{k}$ is almost never feasible, and solving it numerically is expensive. We design an approximate operator which, for an estimate $\hat{\beta}$ of the unknown β in the plant (2) produces an approximate adaptive kernel $\hat{k}$, generated by evaluating a neural operator for the input $\hat{\beta}$.

For such an approach to guarantee stabilization when the exact adaptive kernel $\breve{k}$ is replaced by the approximate kernel $\hat{k}$, we need to design a neural operator that keeps the approximation error $\breve{k}-\hat{k}$ small in a suitable sense. To produce such a neural operator, we recall the DeepONet universal approximation theorem.

Theorem 2 (*DeepOnet Universal Approximation Theorem (Deng et al., 2022)*). *Let $X\subset\mathbb{R}^{d_x}$ and $Y\subset\mathbb{R}^{d_y}$ be compact sets of vectors $x\in X$ and $y\in Y$, respectively. Let $\mathscr{U}: X\to U\subset\mathbb{R}^{d_u}$ and $\mathscr{V}: Y\to V\subset\mathbb{R}^{d_v}$ be sets of continuous functions $u(x)$ and $v(y)$, respectively. Let $\mathscr{U}$ also be compact. Assume the operator $\mathscr{G}:\mathscr{U}\to\mathscr{V}$ is continuous. Then, for all $\epsilon>0$, there exist $m^*, p^*\in\mathbb{N}$ such that for each $m\ge m^*$, $p\ge p^*$, there exist $\theta^{(k)}$, $\vartheta^{(k)}$, neural networks $f^{\mathscr{N}}(\cdot;\theta^{(k)})$, $g^{\mathscr{N}}(\cdot;\vartheta^{(k)})$, $k=1,\dots,p$, and $x_j\in X$, $j=1,\dots,m$, with corresponding $\mathbf{u}_m=(u(x_1),u(x_2),\dots,u(x_m))^T$, such that*

$$|\mathscr{G}(u)(y) - \mathscr{G}_{\mathbb{N}}(\mathbf{u}_m)(y)| < \epsilon, \tag{34}$$

for all functions $u\in\mathscr{U}$ and all values $y\in Y$ of $\mathscr{G}(u)$ where

$$\mathscr{G}_{\mathbb{N}}(y) = \sum_{k=1}^{p} g^{\mathscr{N}}(\mathbf{u}_m;\vartheta^{(k)}) f^{\mathscr{N}}(y;\theta^{(k)}). \tag{35}$$

Note, such a theorem only gives the existence of a neural operator; however, (Bhan et al., 2023), Proposition 1 gives conservative estimates on the number of network parameters needed and further estimates are actively being studied e.g. Mukherjee and Roy (2024).

Let the set H denote the subset of $\mathscr{C}^0([0,1],\mathbb{R})$ endowed with the supremum ($\|\cdot\|_\infty$) norm, such that all $\alpha\in H$ satisfy $\|\alpha\|_\infty\le M$ and α is K-Lipschitz where $M, K>0$ can be as large as required. We denote by $\mathscr{K}: H\to\mathscr{C}^0([0,1],\mathbb{R})$ the operator

$$\mathscr{K}: \hat{\beta}(\cdot,t)\mapsto\breve{k}(\cdot,t), \tag{36}$$

where $\breve{k}(\cdot,t)$ is the solution to the Volterra integral equation (5) at a specific time $t\ge0$. Since the parameter estimate $\hat{\beta}(x,t)$ is time varying, its time derivative affects the closed-loop system. For this reason, it is not enough to approximate only $\breve{k}$. Its derivative $\breve{k}_t(\cdot,t)$ also must be approximated, at each time t. It is crucial to note that, while we are concerned about approximating a derivative in time, $\breve{k}_t(x,t)$, it is only an accurate approximation of this quantity as a function of x that is needed.

For this purpose we denote by $\mathscr{M}: H^2\to\mathscr{C}^0([0,1],\mathbb{R})^2$ the operator

$$\mathscr{M}: (\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t))\mapsto(\mathscr{K}(\hat{\beta}(\cdot,t)),\mathscr{K}_1(\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t))) \tag{37}$$

where $\mathscr{K}_1$ is defined as the operator that maps $(\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t))$ into the solution $\breve{k}_1$ of the Volterra equation

$$\begin{aligned}&\breve{k}_1(x,t) - \int_0^x \hat{\beta}(x-y,t)\breve{k}_1(y,t)dy \\ &+\hat{\beta}_t(x,t) - \int_0^x \hat{\beta}_t(x-y,t)\breve{k}(y,t)dy = 0, \quad x\in[0,1],\end{aligned} \tag{38}$$

namely,

$$\breve{k}_1 = \mathscr{K}_1\left(\hat{\beta},\hat{\beta}_t\right) := \mathscr{B}\left(-\hat{\beta}_t + \hat{\beta}_t * \mathscr{K}\left(\hat{\beta}\right), \mathscr{K}\left(\hat{\beta}\right)\right), \tag{39}$$

which is explicitly given by the expression in Lemma 8. Note, the expression for $\breve{k}_1$ in (38) represents the solution to the time derivative of the kernel $\breve{k}$.

To approximate the operator $\mathscr{M}=(\mathscr{K},\mathscr{K}_1)$ by a DeepONet, the conditions of Theorem 2, require us to define a specific compact set of the $\hat{\beta}(\cdot,t)$, $\hat{\beta}_t(\cdot,t)$ functions. This is the purpose of introducing the set H, which, due to its elements being Lipschitz with a uniform Lipschitz constant, and therefore uniformly equicontinuous, is compact by the Arzelà–Ascoli theorem, and so is the set $H^2 := H\times H$.

We have proven that the operator $\mathscr{M}$ is continuous (and even Lipschitz) in Lemma 2 of Lamarque et al. (2024). We can then state the following theorem, which is a consequence of Theorem 2.

Theorem 3 (*Existence of a Neural Operator Approximating the Kernel*). *For all $\epsilon>0$, there exists a neural operator $(\hat{\mathscr{K}},\hat{\mathscr{K}_1})$ such that for all $\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t)\in H$ and for all $\forall x\in[0,1]$,*

$$\begin{aligned}&\left|\mathscr{K}\left(\hat{\beta}(\cdot,t)\right)(x) - \hat{\mathscr{K}}\left(\hat{\beta}(\cdot,t)\right)(x)\right| \\ &+\left|\mathscr{K}_1\left(\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t)\right)(x) - \hat{\mathscr{K}_1}\left(\hat{\beta}(\cdot,t),\hat{\beta}_t(\cdot,t)\right)(x)\right| < \epsilon.\end{aligned} \tag{40}$$

5. DeepONet-approximated Lyapunov adaptive PDE backstepping design

The stabilizing property of the adaptive backstepping controller employing an approximate estimated kernel is given in the next theorem, our main result.

Theorem 4 (*Stability of Approximate Lyapunov Adaptive Backstepping Control*). *For all $B,c>0$, there exists $\epsilon_0(B,c) = \underset{c\to\infty}{\mathscr{O}}(ce^{\frac{-c}{2}})$, $\gamma_0(B,c) = \underset{c\to\infty}{\mathscr{O}}(e^{-c})$ with a decreasing dependence on* *the argument B such that for all neural operator approximations* $\hat{k} = \hat{\mathscr{K}}\left(\hat{\beta}\right)$ *of accuracy* $\epsilon \in (0, \epsilon_0)$ *provided by Theorem 3, all* $\gamma \in (0, \gamma_0)$*, and all* $\beta, \hat{\beta}(\cdot, 0)$ *that are Lipschitz and satisfy* $\|\beta\|_\infty, \|\hat{\beta}(\cdot, 0)\|_\infty \le B$*, the feedback law*

$$U(t) = \int_0^1 \hat{k}(1-y, t)u(y, t)dy\,, \tag{41}$$

and the update law

$$\hat{\beta}_t(x, t) := \mathrm{Proj}(\tau(x, t), \hat{\beta}(x, t),), \quad \forall (x, t) \in [0, 1] \times \mathbb{R}^+\,, \tag{42}$$

$$\tau(x, t) := \frac{\gamma}{1 + \|w(t)\|_c^2}\Bigg[e^{cx} w(x, t) - \int_x^1 \hat{k}(y - x, t)e^{cy} w(y, t)dy\Bigg]w(0, t), \tag{43}$$

where

$$w = u - \hat{k} * u\,, \tag{44}$$

$$\|w(t)\|_c^2 = \int_0^1 e^{cx} w(x, t)dx\,, \tag{45}$$

guarantee that all solutions for which $\hat{\beta}_t(\cdot, t)$ *remains in H,* $\hat{k}(\cdot, t)$ *is differentiable, and* $\hat{\beta}(\cdot, t)$ *remains Lipschitz for all time satisfy*

$$\Gamma(t) \le R(e^{\rho\Gamma(0)} - 1)\,, \qquad \forall t \ge 0, \tag{46}$$

$$\Gamma(t) = \int_0^1 \left[u^2(x, t) + \left(\beta(x) - \hat{\beta}(x, t)\right)^2\right]dx\,, \tag{47}$$

for constants $\rho, R > 0$ *and, in addition,*

$$u(x, t) \underset{t\to\infty}{\rightarrow} 0\,, \qquad \forall x \in [0, 1]\,. \tag{48}$$

The assumptions that $\hat{\beta}_t(\cdot, t)$ remains in H and that $\hat{k}(\cdot, t)$ is differentiable for all time are strong and not a priori verifiable. They arise from the fact that in the Lyapunov design it is necessary to approximate the update rate $\hat{k}_t$ of the approximated kernel $\hat{k}$. This motivates us to pursue, in Section 7, an alternative *modular design with a passive identifier*, which does not require an approximation of the derivative of the approximated kernel and, hence, does not require these strong assumptions on $\hat{\beta}_t(\cdot, t)$ and $\hat{k}_t(\cdot, t)$.

Our parameter update law (43) is a replica of (14) but with the exact backstepping transformation (4) and the exact kernel (5) replaced, respectively, by the approximate transformation (44) and the DeepONet kernel $\hat{\mathscr{K}}\left(\hat{\beta}\right)$.

We use parameter projection for two reasons. One is for ensuring global stability as in exact adaptive PDE control (Smyshlyaev & Krstic, 2010). The second reason, novel in this paper, is for ensuring that the condition of Theorem 3 remains valid, namely, that $\|\hat{\beta}\|_\infty \le B$ holds for all time. The Lipschitzness of $\hat{\beta}(\cdot, t)$, a technical condition for the Arzela–Ascoli theorem and the compactness of the input set of $\mathscr{K}$, seems impossible to enforce without sacrificing the other more important properties enforced by projection, so we assume it instead.

The elementary pointwise-in-x projection operator (17) has the following well-known properties (Smyshlyaev & Krstic, 2010, Lemma 8.2),

- $\left(\mathrm{Proj}(\tau, \hat{\beta})\right)^2 \le \tau^2, \quad \forall (\tau, \hat{\beta}) \in \mathbb{R} \times [0, B]$ (49)
- If $\hat{\beta}(x, 0) \in [-B, B], \forall x \in [0, 1]$ then the update law $\hat{\beta}_t$ ensures that $\hat{\beta} \in [-B, B]$ (50)
- $-\tilde{\beta}\mathrm{Proj}(\tau, \hat{\beta}) \le -\tilde{\beta}\tau$ for all $\hat{\beta}, \beta \in [-B, B]\,.$ (51)

6. Lyapunov analysis

In this section we prove Theorem 4. We replace the backstepping transformation defined in (4) with its approximate version

$$w(x, t) = u(x, t) - \hat{k} * u(x, t), \quad \forall (x, t) \in [0, 1] \times \mathbb{R}^+\,, \tag{52}$$

obtaining (see Appendix B) the perturbed target system

$$w_t(w, t) = w_x(x, t) + \left[\tilde{\beta}(x, t) - \hat{k} * \tilde{\beta}(x, t)\right]w(0, t) - \Omega(x, t) + w(0, t)\delta(x, t)\,, \tag{53}$$

$$w(1, t) = 0. \tag{54}$$

where

$$\tilde{k} := \check{k} - \hat{k}\,, \tag{55}$$

$$\tilde{\beta} = \beta - \hat{\beta}\,, \tag{56}$$

$$\hat{l} = -\hat{k} + \hat{k} * \hat{l} = \mathscr{K}\left(\hat{k}\right) = \mathscr{K} \circ \hat{\mathscr{K}}\left(\hat{\beta}\right)\,, \tag{57}$$

$$\delta := -\tilde{k} + \hat{\beta} * \tilde{k}\,, \tag{58}$$

$$\Omega = \hat{k}_t * (w - \hat{l} * w)\,. \tag{59}$$

Before commencing our Lyapunov computations, we introduce a lemma on the inverse backstepping kernel $\hat{l}$.

Lemma 2 (*Inverse Kernel Properties*)**.** *Let* $B > 0, \hat{\beta} \in \mathscr{C}^0([0, 1] \times \mathbb{R}^+, \mathbb{R})$ *such that* $\|\hat{\beta}\|_\infty \le B$ *and consider the Volterra equation*

$$\hat{l}(x, t) = -\hat{k}(x, t) + \hat{k} * \hat{l}(x, t), \quad (x, t) \in [0, 1] \times \mathbb{R}^+\,, \tag{60}$$

with the solution $\hat{l} = \mathscr{K} \circ \hat{\mathscr{K}}\left(\hat{\beta}\right)$*, where* $\hat{k} = \hat{\mathscr{K}}\left(\hat{\beta}\right)$ *is defined with* $\hat{\mathscr{K}}$ *provided by Theorem 3 for accuracy* $\epsilon > 0$*. Then*

$$\|\hat{l}\|_\infty \le \bar{k}e^{\bar{k}}\,, \tag{61}$$

$$\bar{k} := Be^B + \epsilon\,. \tag{62}$$

Furthermore, (54) holds if and only if

$$u(x, t) = w(x, t) - \hat{l} * w(x, t)\,, \tag{63}$$

for any pair of functions (u, w)*, and in particular when the state u is governed by (2), (3), and the transformed state* w *is defined by (53), (54).*

Proof. The existence of $\hat{l}$ follows from the facts that it satisfies a Volterra integral equation and that $\hat{k}$ is continuous. The bound (61) is obtained with the successive approximation method, as in the proof of Lemma 1 using (21). To obtain (63), we invoke Lemma 6. □

Lemma 3 (*Lyapunov Estimate for Perturbed Target System*)**.** *For all* $c, B > 0$*, there exist strictly positive quantities* $\epsilon_0(c, B) = \mathscr{O}_{c\to\infty}(ce^{-\frac{c}{2}}), \gamma_0 = \mathscr{O}_{c\to\infty}(e^{-c})$ *with a decreasing dependence on B such that for any* $(\epsilon, \gamma) \in (0, \epsilon_0) \times (0, \gamma_0)$*, any* $\beta, \hat{\beta}(\cdot, 0) \in \mathscr{C}^0([0, 1])$ *that are Lipschitz and satisfy*

$$\|\beta\|_\infty, \|\hat{\beta}(0, \cdot)\|_\infty \le B\,, \tag{64}$$

and for any approximate adaptive backstepping kernel $\hat{k} = \hat{\mathscr{K}}\left(\hat{\beta}\right)$ *provided by Theorem 3 with accuracy* ϵ*, the perturbed target system (53), (54) along with the update law (42) satisfies*

$$|\hat{\beta}(x, t)| \le B, \quad (x, t) \in [0, 1] \times \mathbb{R}^+\,, \tag{65}$$

$$\dot{V}(t) \le -\frac{c}{4}\frac{\|w(t)\|_c^2}{1 + \|w(t)\|_c^2} - \frac{1}{8}\frac{w^2(0, t)}{1 + \|w(t)\|_c^2}\,, \tag{66}$$

where

$$V(t) := \frac{1}{2}\ln\left(1 + \|w(t)\|_c^2\right) + \frac{1}{2\gamma}\int_0^1 \tilde{\beta}^2(x, t)dx\,, \tag{67}$$

$$\|w(t)\|_c^2 := \int_0^1 e^{cx} w^2(x,t)dx\,, \tag{68}$$

for $t \geq 0$.

Proof. The property (65) is immediate, as a result of using projection. With the update law (42), taking the derivative of (67) one arrives at

$$\begin{aligned}
\dot V(t) =\ & \frac{1}{1+\|w(t)\|_c^2}(I_1(t)+I_2(t)+I_3(t)) \\
& + \int_0^1 \underbrace{\frac{w(0,t)e^{cx}\left(1-\hat k*\right)\tilde\beta(x,t)}{1+\|w(t)\|_c^2}}_{=\tau\tilde\beta} dx \\
& + \frac{1}{\gamma}\int_0^1 \underbrace{\left(-\dot{\hat\beta}(x,t)\tilde\beta(x,t)\right)}_{\leq -\tau\tilde\beta \text{ from (42) and (17)}} dx \\
\leq\ & \frac{1}{1+\|w(t)\|_c^2}(I_1(t)+I_2(t)+I_3(t))\,,
\end{aligned} \tag{69}$$

where

$$I_1(t) := w(0,t)\int_0^1 e^{cx}w(x,t)\delta(x,t)\,, \tag{70}$$

$$I_2(t) := -\int_0^1 e^{cx}w(x,t)\Omega(x,t)dx\,, \tag{71}$$

$$I_3(t) := -\frac{1}{2}w^2(0,t) - \frac{c}{2}\|w(t)\|_c^2\,. \tag{72}$$

Using Lemmas 2 and 1, as well as Theorem 3, we have the following upper bounds

$$\|\tilde k\|_\infty \leq \epsilon\,, \tag{73}$$

$$\|\hat k\|_\infty \leq \epsilon + \|\breve k\|_\infty \leq \epsilon + Be^B =: \bar k\,, \tag{74}$$

$$\|l\|_\infty \leq \bar k e^{\bar k} =: \bar l\,, \tag{75}$$

$$\|\delta\|_\infty \leq \epsilon(1+B) =: \bar\delta\epsilon\,, \tag{76}$$

$$\|\hat k_t(t)\| \leq \|\breve k_t(t)\| + \epsilon \leq M\|\hat\beta_t(t)\| + \epsilon\,, \tag{77}$$

where

$$\bar\delta := 1+B\,, \tag{78}$$

$$M := 1+Be^B(2+Be^B)\,, \tag{79}$$

as well as

$$\|\hat\beta_t(t)\| \leq \frac{\gamma}{2}\,\frac{e^{\frac{c}{2}}}{1+\|w(t)\|_c^2}(w^2(0,t)+\|w(t)\|_c^2)(1+\bar k)\,. \tag{80}$$

$I_1(t)$ estimate: Using (76) as well as Young's and Cauchy–Schwarz inequalities, we have the following

$$I_1(t) \leq \frac{w^2(0,t)}{4} + \epsilon^2\bar\delta^2 e^c\|w(t)\|_c^2\,. \tag{81}$$

$I_2(t)$ estimate: We first rework the upper bound (77) using (80)

$$\|\hat k_t(t)\| \leq \frac{\gamma}{2}\times\frac{e^{\frac{c}{2}}w^2(0,t)}{1+\|w(t)\|_c^2}\bar k_t + \frac{\gamma}{2}\times\frac{e^{\frac{c}{2}}\|w(t)\|_c^2}{1+\|w(t)\|_c^2}\bar k_t + \epsilon, \quad t \geq 0\,, \tag{82}$$

$$\bar k_t := M(1+\bar k)\,. \tag{83}$$

Then we use Cauchy–Schwarz inequality to have the following for $t \geq 0$

$$\begin{aligned}
\|\Omega(\cdot,t)\|_\infty \leq\ & \|\breve k_t(t)\|.\|w(t)\|(1+\bar l) + \epsilon\|w(t)\|(1+\bar l) \\
\leq\ & \gamma e^{\frac{c}{2}}\frac{w^2(0,t)}{2}\times\frac{\|w(t)\|_c}{1+\|w(t)\|_c^2}\bar\Omega \\
& + \gamma e^{\frac{c}{2}}\frac{\|w(t)\|}{2}\times\frac{\|w(t)\|_c^2}{1+\|w(t)\|_c^2}\bar\Omega \\
& + \epsilon\|w(t)\|(1+\bar l)\,,
\end{aligned} \tag{84}$$

$$\bar\Omega =: \bar k_t(1+\bar l). \tag{85}$$

With these new inequalities we then have the upper bound for (71) using Cauchy–Schwarz inequality,

$$\begin{aligned}
I_2(t) \leq\ & \gamma e^c\frac{w^2(0,t)}{2}\bar\Omega + \gamma e^c\frac{\|w(t)\|_c^2}{2}\bar\Omega \\
& + \epsilon e^{\frac{c}{2}}\|w(t)\|_c^2(1+\bar l)\,,
\end{aligned} \tag{86}$$

where we have used the fact that $\frac{\|w(t)\|_c^2}{1+\|w(t)\|_c^2} \leq 1$.

Finally, gathering (81), (86), (72) we have that

$$\begin{aligned}
\dot V(t) \leq\ & -\frac{\|w(t)\|_c^2}{1+\|w(t)\|_c^2} \\
& \times\left(\frac{c}{2} - \epsilon^2\bar\delta^2e^c - \gamma e^c\frac{\bar\Omega}{2} - \epsilon e^{\frac{c}{2}}(1+\bar l)\right) \qquad (87)\\
& - \frac{w^2(0,t)}{1+\|w(t)\|_c^2}\left(\frac{1}{4} - \frac{\gamma e^c}{2}\bar\Omega\right). \qquad (88)
\end{aligned}$$

Noting that the quantities $\bar\Omega, \bar l$ depend on ϵ in an increasing fashion, for setting the upper bound on γ, ϵ we fix $\bar\Omega := \bar\Omega(\epsilon = 1)$, $\bar l := \bar l(\epsilon = 1)$. With such fixed choice of $\bar\Omega$ and $\bar l$, all the previous inequalities are valid for all $\epsilon \leq 1$. We now introduce the quantities

$$\gamma_1 := \frac{ce^{-c}}{4\bar\Omega} > 0\,, \tag{89}$$

$$\epsilon_0 := \min\left\{1, \alpha^{-1}\left(\frac{c}{8}\right)\right\} > 0\,, \tag{90}$$

where we introduced the polynomial function $\alpha(\epsilon) = \epsilon^2\bar\delta^2e^c + \epsilon e^{\frac{c}{2}}(1+\bar l)$. Thus, if we choose $\gamma \in (0,\gamma_1)$ and $\epsilon \in (0,\epsilon_0)$ we get that (87) is dominated by $\leq -\frac{c}{4}\frac{\|w(t)\|_c^2}{1+\|w(t)\|_c^2}$. To finish the proof of the lemma, we now consider (88) and introduce the quantity

$$\gamma_0 := \min\left\{\gamma_1, \frac{e^{-c}}{4\bar\Omega}\right\} > 0\,. \tag{91}$$

Taking $\epsilon \in (0,\epsilon_0)$, $\gamma \in (0,\gamma_0)$ gives

$$\dot V(t) \leq -\frac{c}{4}\frac{\|w(t)\|_c^2}{1+\|w(t)\|_c^2} - \frac{1}{8}\frac{w^2(0,t)}{1+\|w(t)\|_c^2}\,. \tag{92}$$

which completes the proof of (66). □

We are now ready to conclude the proof of Theorem 4.

Proof of Theorem 4. Let V be the Lyapunov function defined in (67) and $(\epsilon,\gamma) \in (0,\epsilon_0)\times(0,\gamma_0)$, where ϵ_0, γ_0 are defined in the proof of Lemma 3. It follows from this lemma that $V(t)$ is bounded by $V(0) < \infty$. From the definition of V as (67), we have that $\|\hat\beta\|$, $\|w\|$ are bounded. And by integrating (92) in time and keeping in mind that V is nonnegative, we have the following properties in the sense of norms with respect to time:

- $\|w\| \in \mathscr{L}_2 \cap \mathscr{L}_\infty$
- $w(0,\cdot) \in \mathscr{L}_2$

To achieve the convergence of w to 0, both pointwise and in $\mathscr{L}_2$, and without seeking an H^1 estimate for w, borrowing from the approach in Anfinsen and Aamo (2017, Chapter 4), we introduce the quantity

$$\alpha(x,t) = \mathscr{B}(u(x,t),k(x)), \quad (x,t) \in [0,1]\times\mathbb{R}^+\,, \tag{93}$$

where $k := \mathscr{K}(\beta)$ is the exact backstepping kernel. It follows from (93) that α is a solution to the following transport PDE

$$\alpha_t(x,t) = \alpha_x(x,t), \qquad (x,t) \in [0,1)\times\mathbb{R}^+\,, \tag{94}$$

$$\alpha(1,t) = \int_0^1 (\hat{k}(1-y,t) - k(1-y))u(y,t)dy\,. \tag{95}$$

From Lemma 1, it follows that

$$\|k\|_\infty \leq Be^B. \tag{96}$$

Through the method of characteristics it follows that

$$\alpha(x,t) = \alpha(1, t+x-1), \quad x+t \geq 1\,. \tag{97}$$

and, for $t+x < 1$, we have $\alpha(x,t) = \alpha_0(t+x)$, where α_0 is bounded and denotes the initial condition: $\alpha_0 := u_0 - k * u_0$. We thus have that

$$|\alpha(x,t)| \leq (Be^B + \bar{k})(1+\bar{l})\|w(t+x-1)\|\,, \quad t+x \geq 1\,, \tag{98}$$

and hence $\|\alpha\|_\infty \in \mathscr{L}_\infty$ since we have previously shown that $\|w\| \in \mathscr{L}_\infty$. Since the transformation (93) is invertible,

$$u = \alpha - \beta * \alpha\,, \tag{99}$$

and hence we both have $\|u\|_\infty \in \mathscr{L}_\infty$ and

$$|u(x,t)| \leq (1+B)(Be^B + \bar{k})(1+\bar{l})\|w(t+x-1)\|\,, \quad t+x \geq 1\,. \tag{100}$$

We now prove that $\|w\| \underset{t\to\infty}{\to} 0$ in order to ultimately obtain $\|u\| \underset{t\to\infty}{\to} 0$. Since we already know that $\|w\| \in \mathscr{L}_2$, in order to use Barbalat's lemma, we prove that $\|w\|$ is uniformly continuous by proving that $\frac{d}{dt}\|w\|^2$ is bounded. We first derive the bound

$$\begin{aligned}\left|\frac{d\|w\|^2}{dt}(t)\right| \leq\ & \frac{w^2(0,t)}{2} + 2B\|w(t)\| + 2\bar{k}B\|w(t)\||w(0,t)| \\ & + \gamma e^{\frac{c}{2}}\frac{w^2(0,t)}{2}\bar{\Omega} + \gamma e^{\frac{c}{2}}\frac{\|w(t)\|^2}{2}\bar{\Omega} \\ & + \epsilon\|w(t)\|^2(1+\bar{l}) + \bar{\delta}\epsilon|w(0,t)|\|w(t)\|\,. \end{aligned} \tag{101}$$

Then, recalling that $\|w\| \in \mathscr{L}_\infty \cap \mathscr{L}_2$ and that $w(0,t) = u(0,t)$ is bounded, we have that (101) is bounded. The convergence of $\|w(t)\|$ to zero as time goes to infinity follows from Barbalat's lemma. From (100),

$$\|u(\cdot,t)\|_\infty \underset{t\to\infty}{\to} 0\,. \tag{102}$$

We now prove the global stability (46) in the norm (47). Recalling the Lyapunov functional (67),

$$\|w(t)\|^2 \leq (e^{2V(t)} - 1), \tag{103}$$

$$\|\tilde{\beta}(t)\|^2 \leq 2\gamma V(t) \leq \gamma(e^{2V(t)} - 1), \quad t \geq 0\,. \tag{104}$$

With the inverse backstepping transformation $u = w - \hat{l} * w$ we have the upper bound

$$\|u(t)\|^2 \leq (1+\bar{l})^2\|w(t)\|^2\,. \tag{105}$$

Gathering (105), (103), (104) we have

$$\Gamma(t) \leq \max\left(\gamma, (1+\bar{l})^2\right) \times (e^{2V(t)} - 1)\,. \tag{106}$$

Let us also notice that with the backstepping transformation $w = u - \hat{k} * u$,

$$\frac{1}{2}\ln(1+\|w(t)\|_c^2) \leq \frac{1}{2}e^c\|w(t)\|^2 \leq \frac{1}{2}e^c(1+\bar{k})^2\|u(t)\|^2, \tag{107}$$

which leads to

$$2V(t) \leq \max\left(\frac{1}{\gamma}, e^c(1+\bar{k})^2\right) \times \Gamma(t), \quad t \geq 0. \tag{108}$$

Gathering (108) and (106) we have the following

$$\Gamma(t) \leq R(e^{\rho\Gamma(0)} - 1), \quad t \geq 0\,, \tag{109}$$

$$R := \max(\gamma, (1+\bar{l})^2)\,, \tag{110}$$

$$\rho := \max\left(\frac{1}{\gamma}, e^c(1+\bar{k})^2\right)\,. \tag{111}$$

Note that the coefficients R, ρ depend in an increasing fashion on ϵ. To make them independent of the approximation accuracy ϵ, one can choose $\rho := \rho(\epsilon = 1), R := R(\epsilon = 1)$ and all the results are still valid as long as we train the DeepONet $\hat{\mathscr{K}}$ for $\epsilon \in (0, \min(1, \epsilon_0))$. Further, note that ϵ_0 explicitly depends on the max size of the family of β functions given by B and thus a smaller B will enable larger neural operator approximation error. □

7. A modular design with a passive identifier

In this section we depart from the Lyapunov adaptive design of the previous sections and employ a *passive identifier* design instead. For ODEs, this identifier is introduced in Krstic, Kanellakopoulos, and Kokotovic (1995, Chapter 5). Its first use in adaptive control of PDEs is in Smyshlyaev and Krstic (2007a, Sections implying equicontinuity and is uniformly bounded by2.1, 3, and 4), for parabolic PDEs. The first use of a passive identifier in control of a hyperbolic PDE is in Anfinsen and Aamo (2019, Chapter 4).

Compared to the Lyapunov design, in which the states of the entire system (the plant and the parameter estimator) are captured in a single Lyapunov function, the passive identifier design neither offers superior performance nor the lowest possible dynamic order. In fact, its dynamic order is increased due to the redundancy of the measured state $u(x,t)$ being estimated by another PDE observer state, $\hat{u}(x,t)$, whose sole role is in the estimation of the unknown parameter. However, the reward for using this less dynamically efficient approach is that the conditions for the estimation of the gain kernel operator are less stringent and the analysis is freed of the requirement to estimate the time derivative of the approximate kernel.

We start by introducing a passive observer-based identifier. For the u-system (2), (3), linearly parametrized in the functional coefficient $\beta(x)$, as proposed in Anfinsen and Aamo (2019, (4.5)), we introduce the observer

$$\hat{u}_t(x,t) = \hat{u}_x(x,t) + \hat{\beta}(x,t)u(0,t) + \gamma_0(u(x,t) - \hat{u}(x,t))u^2(0,t) \tag{112}$$

$$\hat{u}(1,t) = U(t)\,. \tag{113}$$

where $\gamma_0 > 0$ and the term $\gamma_0(u(x,t) - \hat{u}(x,t))u^2(0,t)$ represents a form of nonlinear damping in the observer, which plays the same role as update law normalization (namely, to bound the parameter update rate, $\hat{\beta}_t(x,t)$), and which was introduced in the x-passive scheme in Smyshlyaev and Krstic (2007a, Section 5.6).

For the parameter update law, we employ a slight modification of Anfinsen and Aamo (2019, (4.6)),

$$\tau(x,t) := \gamma(u(x,t) - \hat{u}(x,t))u(0,t)\,, \tag{114}$$

$$\hat{\beta}_t(x,t) := \text{Proj}\,(\tau(x,t), \tau_x(x,t))\,, \quad (x,t) \in [0,1] \times \mathbb{R}^+\,, \tag{115}$$

where $\gamma > 0$ and the Proj is defined in (17). From Anfinsen and Aamo (2019, Lemma 4.1), we get the following result, in which the spaces $\mathscr{L}_2$ and $\mathscr{L}_\infty$ are with respect to $t \in [0, \infty)$.

Lemma 4 (*Properties of Passive Identifier (Anfinsen & Aamo, 2019, Lemma 4.1)*). *The identifier (112)–(113), with an arbitrary initial condition $\hat{u}_0 = \hat{u}(\cdot,0)$ such that $\|\hat{u}_0\| < \infty$, along with the update law (115) with an arbitrary Lipschitz initial condition $\hat{\beta}_0 = \hat{\beta}(\cdot,0)$ such that $\|\hat{\beta}_0\|_\infty \leq B$, guarantees that all solutions satisfy*

$$\|\hat{\beta}(\cdot,t)\| \leq B, \quad t \geq 0\,, \tag{116}$$

$$\|e\| \in \mathscr{L}_\infty \cap \mathscr{L}_2\,, \tag{117}$$

$$|e(0,\cdot)|,\ \|e\||u(0,\cdot)|,\ \|\hat{\beta}_t\| \in \mathscr{L}_2\,, \tag{118}$$

where

$$e := u - \hat{u}\,. \tag{119}$$

Next, we introduce our adaptive control law with a DeepONet-approximated gain. Let us first recall the definition of the exact estimated kernel $\breve{k} = \mathscr{K}(\hat{\beta})$ through the solution of the Volterra equation

$$\breve{k}(x,t) = -\hat{\beta}(x,t) + \int_0^x \hat{\beta}(y,t)\breve{k}(x-y,t)dy, \tag{120}$$

for all $(x,t) \in [0,1] \times \mathbb{R}^+$. A weaker version of an approximating operator $\hat{\mathscr{K}}$ for the approximate estimator kernel $\hat{k} = \hat{\mathscr{K}}(\hat{\beta})$ suffices as compared to the approximation in Theorem 3.

Theorem 5 (*Existence of a NO to Approx. the Kernel*). *For all $\epsilon > 0$ there exists a neural operator $\hat{\mathscr{K}}$ such that for all $\hat{\beta}(\cdot,t) \in H$, for all $\forall x \in [0,1]$,*

$$\left|\mathscr{K}\left(\hat{\beta}(\cdot,t)\right)(x) - \hat{\mathscr{K}}\left(\hat{\beta}(\cdot,t)\right)(x)\right| < \epsilon\,. \tag{121}$$

We are now ready to state an equivalent of Theorem 4

Theorem 6 (*Stability of Approximate Passive-Identifier Adaptive Backstepping Control*). *For all $B, \gamma, \gamma_0 > 0$ and $\epsilon_0 := \frac{e^{-\frac{3}{2}}}{\sqrt{2}(1+B)} > 0$ such that for all neural operator approximations $\hat{k}$ of accuracy $\epsilon \in (0,\epsilon_0)$ provided by Theorem 5, the plant (2),(3), in feedback with the adaptive control law*

$$U(t) = \int_0^1 \hat{k}(x-y,t)\hat{u}(y,t)dy\,, \tag{122}$$

along with the update law for $\hat{\beta}$ given by (115) with any Lipschitz initial condition $\hat{\beta}_0 = \hat{\beta}(\cdot,0)$ such that $\|\hat{\beta}_0\|_\infty \le B$, and the passive observer $\hat{u}$ given by (112), (113) with any initial condition $\hat{u}_0 = \hat{u}(\cdot,0)$ such that $\|\hat{u}_0\| < \infty$, satisfies the following properties for all solutions for which $\hat{\beta}(\cdot,t)$ remains Lipschitz for all time:

$$\|u\|,\ \|\hat{u}\|,\ \|u\|_\infty,\ \|\hat{u}\|_\infty \in \mathscr{L}_2 \cap \mathscr{L}_\infty\,, \tag{123}$$

$$\|u\|_\infty,\ \|\hat{u}\|_\infty \to 0\,, \quad \text{as } t \to \infty\,. \tag{124}$$

Additionally, the following global stability estimate holds for the equilibrium $(u,\hat{u},\hat{\beta}) = (0,0,\beta)$,

$$S(t) \le RS(0)e^{\rho S(0)}, \quad t \ge 0, \tag{125}$$

where

$$S := \|u\|^2 + \|\hat{u}\|^2 + \|\tilde{\beta}\|^2\,, \tag{126}$$

and $\rho, R > 0$ are strictly positive constants.

Proof. The proof borrows considerably from Anfinsen and Aamo (2019, Chapter 4), with minimum repetition, and with necessary augmentation to account for the gain approximation error $\breve{k} - \hat{k}$.

Part A: Perturbed target system. We take the same *exact* adaptive backstepping transformation as (4) but apply it to the observer state $\hat{u}$, namely,

$$w(x,t) := \hat{u}(x,t) - \int_0^x \breve{k}(x-y,t)\hat{u}(y,t)dy, \quad (x,t) \in [0,1] \times \mathbb{R}^+\,, \tag{127}$$

where $\breve{k}$ is the exact solution to the Volterra equation (5). This backstepping transformation leads to the following system satisfied by w (for the computations refer to Appendix C):

$$w_t(x,t) = w_x(x,t) - \breve{k}(x,t)e(0,t) + \gamma_0 u^2(0,t)\mathscr{B}(e,\breve{k})(x,t) + \Omega(x,t)\,, \tag{128}$$

$$w(1,t) = -\int_0^1 \tilde{k}(1-y,t)\mathscr{B}(w,\hat{\beta})(y,t)dy =: \Gamma(t), \quad (x,t) \in [0,1] \times \mathbb{R}^+\,, \tag{129}$$

where

$$\mathscr{B}(e,\breve{k}) := e - \breve{k} * e\,, \tag{130}$$

$$\mathscr{B}(w,\hat{\beta}) := w - \hat{\beta} * w\,, \tag{131}$$

$$\Omega(x,t) := \int_0^x \breve{k}_t(x-y,t)\mathscr{B}(w,\hat{\beta})(y,t)dy\,, \tag{132}$$

$$\tilde{k} := \breve{k} - \hat{k}\,. \tag{133}$$

Notice that the only difference with the system described in Anfinsen and Aamo (2019, (4.29)) lies in the presence of perturbed boundary conditions Γ, which is a consequence of the controller choice U that employs an approximated estimated kernel $\hat{k}$ instead of the exact estimated kernel $\breve{k}$.

Spatial $\mathscr{L}^2$ boundedness and regulation of plant and observer states. We use the following Lyapunov function candidate (Anfinsen & Aamo, 2019, (4.42)):

$$V(t) := \|w(t)\|_c^2 = \int_0^1 e^{cx} w^2(x,t)dx, \quad t \ge 0\,, \tag{134}$$

where $c > 0$ is an arbitrary positive constant. Before starting the Lyapunov computations we first state and recall inequalities that can be achieved from Lemma 1

$$\|\breve{k}\|_\infty \le Be^B := \bar{k}\,, \tag{135}$$

$$\|\tilde{k}\|_\infty \le \epsilon\,, \tag{136}$$

$$\|\breve{k}_t(t)\| \le M\|\hat{\beta}_t(t)\|\,, \tag{137}$$

$$|\Gamma(t)| \le \epsilon\bar{\Gamma}\|w(t)\|\,, \tag{138}$$

$$\|w(t)\| \le G_1\|\hat{u}(t)\|\,, \tag{139}$$

$$\|\hat{u}(t)\| \le G_2\|w(t)\|\,, \tag{140}$$

where

$$M := 1 + Be^B(2 + Be^B)\,, \tag{141}$$

$$\bar{\Gamma} := 1 + B\,, \tag{142}$$

$$G_1 := 1 + \bar{k}\,, \tag{143}$$

$$G_2 := 1 + B\,. \tag{144}$$

We use the same computations as the one done in Anfinsen and Aamo (2019, Chapter 4) with the only difference that $w^2(1,t) = \Gamma^2(t) \neq 0$. We also choose $c = 3$ and it leads to the following upper bound

$$\dot{V}(t) \le -V(t)\left(1 - e^c\epsilon^2\bar{\Gamma}^2\right) + l_1(l)V(t) + l_2(t), \quad t \ge 0\,, \tag{145}$$

where

$$l_1(t) := 2G_1^2\gamma_0^2 e^{2c}\|e(t)\|^2u^2(0,t) + e^cG_2^2\|\breve{k}_t\|^2\,, \tag{146}$$

$$l_2(t) := (e^3\bar{k}^2 + 1)e^2(0,t), \quad t \ge 0\,. \tag{147}$$

We introduce

$$\epsilon_0 := \frac{e^{-\frac{c}{2}}}{\sqrt{2}\bar{\Gamma}}\,. \tag{148}$$

Thus, if we choose $\epsilon \in (0,\epsilon_0)$ we have $1 - e^c\epsilon^2\bar{\Gamma}^2 > \frac{1}{2} > 0$. Since, from (137), $\|\breve{k}_t\| \le M\|\hat{\beta}_t\|$, we have from Lemma 4 that $l_1, l_2 \in \mathscr{L}^1$ (and are positive). Then using Krstic et al. (1995, Lemma B.6) we have that

$$V \in \mathscr{L}_1 \cap \mathscr{L}_\infty, \quad V(t) \underset{t\to\infty}{\to} 0. \tag{149}$$

It follows from (149) that $\|w\| \in \mathscr{L}_2 \cap \mathscr{L}_\infty$, $\|w(t)\| \underset{t\to\infty}{\to} 0$. Further, from (140) we have the same for $\hat{u}$. Lastly, from Lemma 4 it follows that $\|u\| \in \mathscr{L}_2 \cap \mathscr{L}_\infty$.

Part B: Pointwise-in-space boundedness and regulation. Exactly like in Anfinsen and Aamo (2019, (3.11)) we also introduce the quantity

$$\alpha(x,t) = u(x,t) - \int_0^x k(x-y)u(y,t)dy\,, \tag{150}$$

for all $(x,t) \in [0,1] \times \mathbb{R}^+$ with k being the exact backstepping kernel i.e $k = \mathscr{K}(\beta)$. The backstepping transformation of (150) leads to the following transport PDE

$$\alpha_t = \alpha_x\,, \tag{151}$$

$$\alpha(1,t) = \int_0^1 \hat{k}(1-y,t)\hat{u}(y,t)dy - \int_0^1 k(1-y)u(y,t)dy, \quad (x,t) \in [0,1]\times\mathbb{R}^+\,. \tag{152}$$

The only difference with Anfinsen and Aamo (2019, (4.57b)) lies in the presence $\hat{k}$ instead of $\check{k}$ in the boundary condition (152). But noticing that thanks to (136), (135) we have

$$|\hat{k}(x,t)| \le \epsilon + \bar{k}\,. \tag{153}$$

Thus $\alpha(1,t)$ remains bounded. The solution of the transport PDE (151)–(152) is given by

$$\alpha(x,t) = \alpha(1, t+x-1), \quad x+t \ge 1\,. \tag{154}$$

and, for $t+x < 1$, we have $\alpha(x,t) = \alpha_0(t+x)$, where α_0 is bounded and denotes the initial condition: $\alpha_0 := u_0 - k * u_0$. It follows that $\|\alpha\|_\infty \in \mathscr{L}_\infty$. Since the transformation (150) is invertible, $u = \alpha - \beta * \alpha$, we also have that $\|u\|_\infty \in \mathscr{L}_\infty$. We then achieve an upper bound on $\frac{d}{dt}\|u\|^2$ to get the regulation to 0 of $\|u\|$ through Barbalat's lemma. From

$$\left|\frac{d\|u\|^2}{dt}(t)\right| \le U^2(t) + u^2(0,t) + 2B|u(0,t)|\|u(t)\| < \infty\,, \tag{155}$$

we have that $\|u(t)\| \underset{t\to\infty}{\to} 0$. Since $\|u(t)\|, \|\hat{u}(t)\| \to 0$, we also have $\alpha(1,t) \to 0$. From the last observation it follows that

$$\|\alpha(\cdot,t)\|_\infty \underset{t\to\infty}{\to} 0, \quad t \mapsto \|\alpha(\cdot,t)\|_\infty \in \mathscr{L}_2 \cap \mathscr{L}_\infty\,. \tag{156}$$

With the invertibility of the transformation (150), namely, $u = \alpha - \beta * \alpha$, we have that

$$\|u\|_\infty \in \mathscr{L}_\infty \cap \mathscr{L}_2, \qquad \|u(t)\|_\infty \underset{t\to\infty}{\to} 0. \tag{157}$$

We now prove a similar result for $\hat{u}$. To do so we first use the change of variable

$$\hat{e}(x,t) := e(1-x,t). \tag{158}$$

This leads to the following PDE satisfied by $\hat{e}$

$$\hat{e}_t(x,t) + \hat{e}_x(x,t) = a(t)\hat{e}(x,t) + f(x,t), \tag{159}$$

$$\hat{e}(0,t) = 0\,, \tag{160}$$

where

$$f(x,t) = \tilde{\beta}(1-x,t)u(0,t)\,, \tag{161}$$

$$a(t) = -\gamma_0 u^2(0,t). \tag{162}$$

We are now ready to use Karafyllis and Krstic (2020, Theorem 2.3) to achieve the following ISS result for $\hat{e}$ for $t \ge 1$

$$\begin{aligned}\|\hat{e}(\cdot,t)\|_\infty &\le 2Be^{\left(1+\mu-\gamma_0 \min_{0\le s\le t} u^2(0,s)\right)} \max_{t-1\le s\le t}\left(|u(0,t)|e^{-\mu(t-s)}\right)\\ &\le 2Be^{1+\mu}\max_{t-1\le s\le t}|u(0,t)|\,,\end{aligned} \tag{163}$$

where

$$\mu := 2\gamma_0 \max_{t\ge 0} u^2(0,t) < \infty\,, \tag{164}$$

since $\|u\|_\infty \in \mathscr{L}_\infty$. From (163) we are now ready to prove that $\|\hat{e}\|_\infty \in \mathscr{L}_2 \cap \mathscr{L}_\infty$, $\|\hat{e}\|_\infty \underset{t\to\infty}{\to} 0$. Notice that from (150) we have that $u(0,s) = \alpha(0,s)$. From (154) we thus have for $t \le s \le t+1$

$$|u(0,s)| = |\alpha(0,s)| = |\alpha(1,s-1)| = |\alpha(s-t,t)|\,, \tag{165}$$

$$\max_{t\le s\le t+1}|u(0,s)| \le \|\alpha(\cdot,t)\|_\infty\,. \tag{166}$$

Since $\|\alpha\|_\infty \in \mathscr{L}_2$, $\|\alpha(t)\|_\infty \underset{t\to\infty}{\to} 0$, from (163), (166) the same holds for $\hat{e}$, and thus for e and the same for $\hat{u}$ since $\hat{u} = u - e$.

Part C: Global stability. We now prove (125). Define

$$S(t) := \|u(t)\|^2 + \|\hat{u}(t)\|^2 + \|\tilde{\beta}(t)\|^2, \quad t \ge 0\,. \tag{167}$$

The goal is to prove the existence of a function $\theta \in \mathscr{K}_\infty$ such that

$$S(t) \le \theta(S(0)), \quad t \ge 0\,. \tag{168}$$

We begin by reusing the Lyapunov functions introduced in Anfinsen and Aamo (2017, Chapter 4)

$$V_1(t) := \int_0^1 (1+x)\left[e^2(x,t) + \frac{1}{\gamma}\tilde{\beta}^2(x,t)dx\right], \quad t \ge 0\,. \tag{169}$$

Using the proof of Anfinsen and Aamo (2019, Lemma 4.1) leads to the following

$$\int_0^\infty e^2(0,\tau)d\tau + \int_0^\infty \|e(\tau)\|^2 d\tau + 2\gamma_0 \int_0^\infty \|e(\tau)\|^2 u^2(0,\tau)d\tau \le V_1(0)\,. \tag{170}$$

Also from the definition of the update law (115) we have that

$$\|\check{k}_t(t)\|^2 \le M^2\|\hat{\beta}_t(t)\|^2 \le \frac{M^2\gamma^2}{2\gamma_0}(2\gamma_0\|e(t)\|^2 u^2(0,t))\,. \tag{171}$$

Recalling (145), we also have from Krstic et al. (1995, Lemma B.6) that

$$V(t) \le (e^{-\frac{t}{2}}V(0) + \|l_2\|_1)e^{\|l_1\|_1}\,. \tag{172}$$

Then recalling (146), (147) (171) and (170) we have

$$\|l_1\|_1 \le \bar{l}_1 V_1(0)\,, \tag{173}$$

$$\|l_2\|_2 \le \bar{l}_2 V_1(0)\,, \tag{174}$$

where, recalling that $\bar{k} = Be^B, G_1 = Be^B, G_2 = B, M(B) = 1 + Be^B(2+Be^B)$, the coefficients $\bar{l}_1, \bar{l}_2$ are given by

$$\bar{l}_1(B,\gamma,\gamma_0) := \max\left(\left(1+Be^B\right)^2\gamma_0 e^{2c}, \frac{\gamma^2 e^c(1+B)^2\left(1+Be^B(2+Be^B)\right)^2}{2\gamma_0}\right), \tag{175}$$

$$\bar{l}_2(B) := 1 + e^3\left(1+Be^B\right)^2\,. \tag{176}$$

We then introduce the function

$$V_3(t) := V_1(t) + V(t) = \int_0^1 (1+x)\left[\frac{\tilde{\beta}(x,t)^2}{\gamma} + e^2(x,t)\right] + \int_0^1 e^{3x}w^2(x,t)dx\,. \tag{177}$$

Noticing that

$$V_1(t) \le \bar{l}_2 V_1(0)e^{\|l_1\|_1}, \quad t \ge 0\,, \tag{178}$$

we achieve from (172), (178), (173) and (174) the following

$$V_3(t) \leq 2\bar{l}_2 V_3(0) e^{\bar{l}_1 V_3(0)}, \quad t \geq 0. \tag{179}$$

We thus have the following for $t \geq 0$ using the Cauchy–Schwarz and Young inequalities

$$\begin{aligned} V_3(t) &\geq \frac{1}{\gamma}\|\tilde{\beta}(t)\|^2 + \|e(t)\|^2 + \frac{\|\hat{u}(t)\|^2}{(1+B)^2} \\ &\geq \frac{1}{\gamma}\|\tilde{\beta}(t)\|^2 + \frac{1}{(1+B)^2}(\|e(t)\|^2 + \|\hat{u}(t)\|^2) \\ &\geq \frac{1}{\gamma}\|\tilde{\beta}(t)\|^2 + \frac{1}{(1+B)^2}(\|u(t)\|^2 + 2\|\hat{u}(t)\|^2 \\ &\quad - 2\|u(t)\|\|\hat{u}(t)\|) \\ &\geq \frac{1}{\gamma}\|\tilde{\beta}(t)\|^2 + \frac{1}{(1+B)^2}\left(\frac{1}{4}\|\hat{u}(t)\|^2 + \frac{1}{3}\|u(t)\|^2\right) \\ &\geq \min\left(\frac{1}{\gamma}, \frac{1}{4(1+B)^2}\right) S(t). \end{aligned} \tag{180}$$

We now focus on establishing the upper bound on V_3. From (177) we have for $t \geq 0$ with Young's inequality

$$\begin{aligned} V_3(t) &\leq \frac{2}{\gamma}\|\tilde{\beta}(t)\|^2 + 4\|u(t)\|^2 + 4\|\hat{u}(t)\|^2 + e^3(1+\bar{k})^2\|\hat{u}(t)\|^2 \\ &\leq \max\left(\frac{2}{\gamma}, 4 + e^3(1+\bar{k})^2\right) S(t). \end{aligned} \tag{181}$$

Then gathering (179), (180), (181) we obtain (125) with

$$R(B, \gamma, \gamma_0) := 2\bar{l}_2 \max\left(\gamma, 4(1+B)^2\right), \tag{182}$$

$$\rho(B, \gamma, \gamma_0) := \bar{l}_1 \max\left(\frac{2}{\gamma}, 4 + e^3\left(1 + Be^B\right)^2\right). \quad \square \tag{183}$$

Examining the bounds R and ρ in (182), (183), in light of (175) and (176), one notes their explicit, albeit conservative dependence on the "instability bound" B, the adaptation gain γ, and the normalization (observer nonlinear damping) gain γ_0. The increasing dependence on the instability B is the most evident, and expected.

8. Simulations

We simulate the system governed by (2), (3) where the plant coefficient $\beta(x) = 5\cos(\sigma \cos^{-1}(x))$ is defined as a Chebyshev polynomial with shape parameter σ. This choice of $\beta(x)$ follows from Bhan et al. (2023), Lamarque et al. (2024), as they are a dense orthogonal L^2 set of functions and therefore approximate a large number of expected functions in practice. However, we emphasize that any compact set of continuous functions can be chosen for the plant coefficients $\beta(x)$ (e.g. Fourier series). For simulation of the hyperbolic PDE, we utilize a first-order upwind scheme with temporal step $dt = 5 \times 10^{-4}$ and spatial step $dx = 1 \times 10^{-2}$. We note that the given PDE with $\beta(x)$ as a Chebyshev polynomial is open-loop unstable (Figure 3, Bhan et al. (2023)). For the adaptive control scheme, we utilize the Lyapunov approach given in (41), (42), (43), (44) with a first order Euler scheme for (42).

We now begin our discussion on training the NO-approximated kernel. To effectively handle the adaptive estimates of $\hat{\beta}(x)$ and the corresponding kernels, one must construct a diverse and exhaustive dataset anticipating the possible β functions encountered. The simplest way to build this dataset is by generating β values with varying σ and simulating the true adaptive controller saving both the β functions and corresponding kernels encountered. Although simulating the true adaptive controller is expensive, the construction of the dataset, like training, only needs to be done once offline. In this work, we considered 10 β functions with $\sigma \sim$ Uniform(2.7, 3.2) and simulate the resulting PDEs under the adaptive controller for $T = 10$s, sub-sampling each pair of (β, k) every 0.01s using the finite difference scheme in Bhan, Bian, Krstic, and Shi (2024) for the kernel calculation k. This creates a total dataset of 10000 different (β, k) pairs to perform supervised learning of the neural operator (available publicly https://t.ly/w2kFR). We note that if one wants to handle a larger family of plant coefficients, they will need to sample more β functions and perform similar calculations running the true adaptive controller. Lastly, we briefly mention that although the Lyapunov approach as discussed in Sections 4, 5 requires approximation of the derivatives, we found sufficient performance without the calculation intensive derivative approximation (See Lamarque et al. (2024, Sec. XI) for more details on neural operator approximation of derivatives).

Table 2
Neural operator speedups over the analytical kernel calculation with respect to the increase in discretization points (decrease in step size).

Spatial step size (dx)	Analytical kernel calculation time (s) ↓	Neural operator kernel calculation time (s) ↓	Speedup ↑
0.01	0.044	0.023	1.87x
0.001	2.697	0.024	110x
0.0005	10.334	0.024	427x
0.0001	245	0.037	6642x

The training of the NO uses the DeepXDE package (Lu, Meng, Mao, & Karniadakis, 2021) and requires approximately 100 seconds to train (whereas the dataset takes several minutes to construct). The resulting DeepONet consists of 14913 parameters with traditional multi-layer perceptron (MLPs) for the branch and trunk networks. Despite the small network, excellent accuracy is achieved as the L_2 training error was 2×10^{-3} and the L_2 testing error was 1.8×10^{-3}.

We begin our discussion of the numerical simulations by presenting NO speedups averaged over 100 calculations of the kernel, according to discretization size, in Table 2. We can see that as the spatial step size grows, the speedup increases shrinking the analytical kernel calculation time from *4 min to 0.4 seconds*. This is only for a single kernel calculation in which the speedup is exemplified as in each timestep in adaptive control, the resulting kernel needs to be continually recalculated according to the new $\hat{\beta}$ estimate.

Lastly, we conclude by presenting a single instance of the resulting controller under NO approximated kernels in Fig. 1. This instance presents $\beta(x)$ as the aforementioned Chebyshev polynomial with $\sigma = 2.9$ and initializes the estimated plant parameter to $\hat{\beta}(x, 0) = 1$. We emphasize that this specific $\beta(x)$ was not seen in any of the $\beta(x)$ functions utilized for training. In Fig. 1, the plant's instability in the first eight seconds drives the estimation of $\hat{\beta}$, but then, by ten seconds, the estimate is good enough to provide a stabilizing controller leading to rapid decay of the system state. Furthermore, the stabilization annihilates the persistence of excitation from the plant's estimator leading to the stagnation of the estimate $\hat{\beta}$. This is observed clearly in Fig. 2 where $\hat{\beta}$ freezes by $t = 10$ and — due to lack of excitation — never reaches the true $\beta(x)$ value. We stress that this lack of convergence towards the true β is not an issue but merely a feature of adaptive control as one is not performing perfect plant identification, but estimating with the goal of stabilization, which is aptly achieved with the final, inexact $\hat{\beta}$ (in red) of Fig. 2. We conclude our discussion with Fig. 3 showcasing the kernel computed using the NO over time. As expected, once $\hat{\beta}$ stalls, the corresponding kernel — which is a mapping relying solely on $\hat{\beta}$ — stagnates concurrently. Furthermore, in the right of Fig. 3 we see that the NO approximation is very close to the analytical estimate maximizing at a relative error of approximately 10%.

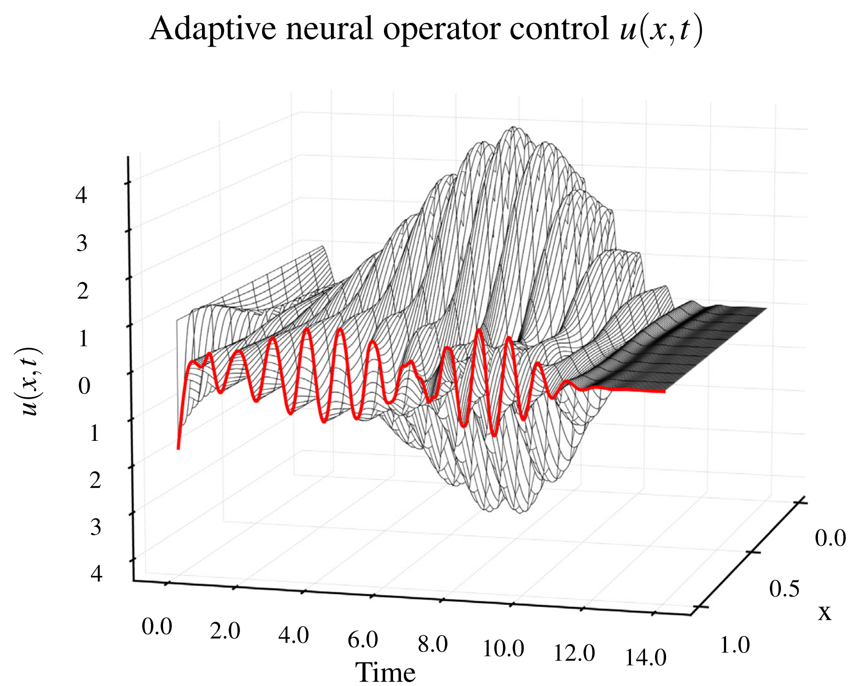


Fig. 1. Adaptive neural operator controller applied to the PDE governed by (2), (3) where $\beta(x) = 5\cos(\sigma \cos^{-1}(x))$ with $\sigma = 2.9$ and initial condition $u(x, 0) = 1$. The initial guess for $\hat{\beta}$ was $\hat{\beta}(x, 0) = 1 \quad \forall x \in [0, 1]$ and the control update law (42), (43), (44), (45) has parameter $c = 1$.

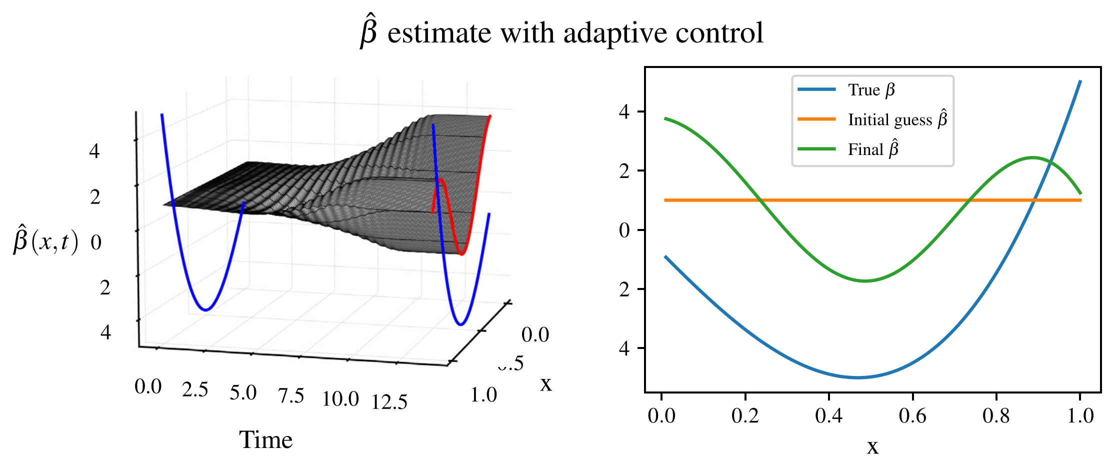


Fig. 2. Left: $\hat{\beta}$ estimates when controlling the PDE in Fig. 1 using neural operator approximated kernels: true β (blue) and final estimated $\hat{\beta}$ (red); Right: comparison between the true β value, the initial guess $\hat{\beta}(\cdot, 0) = 1$, and the final estimated $\hat{\beta}$ at $t = 13$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

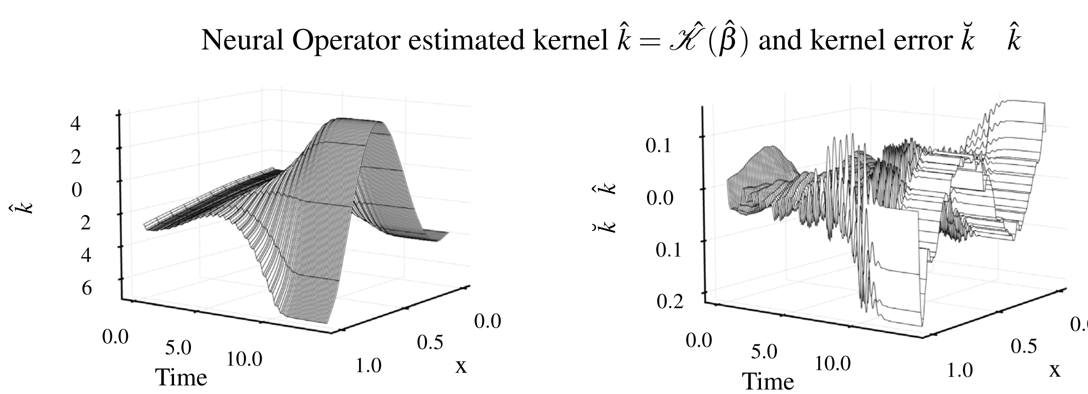


Fig. 3. Neural operator approximated kernels when controlling the PDE in Fig. 1 (left), and the difference in kernel error between the approximated kernel and the analytical kernel (right).

9. Conclusion

In this paper, we present the first results for NO approximated kernels in adaptive control of hyperbolic PDEs. We consider two approaches, namely a Lyapunov-based approach and a modular approach with a passive identifier, and prove global stability for both approaches, with tradeoffs between assumptions and dynamic orders. We then present numerical simulations showcasing the viability of the Lyapunov approach under the neural operator approximated kernels obtaining speedups on the magnitude of 10^3. With such large reduction in computational costs, NO-based adaptive backstepping opens the door for applying adaptive PDE control in real-time.

Acknowledgments

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

Appendix A. Backstepping transformation and involution operator for the kernel

We introduce a *backstepping operator* $\mathscr{B}$ defined as

$$\mathscr{B}(\xi,\eta) := \xi - \eta * \xi\,, \tag{A.1}$$

and, with this operator, introduce the (Volterra-type) *backstepping equation*

$$\mathscr{B}(\xi,\eta) = \zeta\,, \tag{A.2}$$

meant to be solved for ξ, for given (ζ,η). We denote the solution of (A.2) for ξ with the operator $\mathscr{W}(\zeta,\eta)$. Next, setting $\zeta=-\eta$ in (A.2), we introduce the *kernel integral equation*

$$\mathscr{B}(\xi,\eta)+\eta = \xi - \eta*\xi+\eta = 0\,, \tag{A.3}$$

and denote its solution for ξ with the operator $\mathscr{K}(\eta) := \mathscr{W}(-\eta,\eta)$, namely, as

$$\mathscr{B}(\mathscr{K}(\eta),\eta)+\eta = \mathscr{K}(\eta)-\eta*\mathscr{K}(\eta)+\eta = 0\,. \tag{A.4}$$

Next, we give a previously unobserved property of $\mathscr{K}$.

Lemma 5. *$\mathscr{K}^{-1}=\mathscr{K}$, i.e. $\mathscr{K}^2 := \mathscr{K}\circ\mathscr{K} = Id$.*

Proof. By noting that the roles of ξ and η in (A.3) are interchangeable, or by using the Laplace transform. □

Due to the property given by Lemma 5, we call $\mathscr{K}$ the *involution operator*.[2]

The next lemma gives an explicit expression for the operator $\mathscr{W}$.

Lemma 6.

$$\begin{aligned}\mathscr{W}(\zeta,\eta) &= \zeta - \mathscr{K}(\eta)*\zeta \\ &= \mathscr{B}(\zeta,\mathscr{K}(\eta)) \\ &= \mathscr{B}\left(\zeta,\mathscr{K}^{-1}(\eta)\right)\,.\end{aligned} \tag{A.5}$$

Proof. By direct substitution into (A.2), or by using the Laplace transform. The last equality follows from Lemma 5. □

[2] Because a matrix A such that $A^2=I$ is typically referred to as *involutory*.

To summarize,

$$\zeta = \mathscr{B}(\xi,\eta) \quad \text{iff} \quad \xi = \mathscr{B}\left(\zeta,\mathscr{K}^{-1}(\eta)\right)\,, \tag{A.6}$$

or, alternatively stated, if $\xi+\eta=\eta*\xi$, then

$$w = u-\xi*u \quad \text{iff} \quad u = w-\eta*w\,. \tag{A.7}$$

These observations yield the following result.

Lemma 7. *The operator $(\eta,\zeta)\mapsto(\mathscr{K}(\eta),\mathscr{B}(\zeta,\eta))$ is an involution.*

Proof. By noting that

$$\mathscr{K}(\mathscr{K}(\eta)) = \eta \tag{A.8}$$

$$\mathscr{B}\left(\mathscr{B}(\zeta,\eta),\mathscr{K}(\eta)\right) = \zeta\,. \quad \square \tag{A.9}$$

In calculations to come, Eq. (A.2) will arise in a particular form. We provide its solution in the following lemma.

Lemma 8. *For given functions β_0, β_1, and $k_0=\mathscr{K}(\beta_0)$, if the function k_1 satisfies the equation*

$$k_1-\beta_0*k_1+\beta_1-\beta_1*k_0 = 0\,, \tag{A.10}$$

then it is explicitly given by

$$\begin{aligned}k_1 = \mathscr{K}_1(\beta_0,\beta_1) := &-\beta_1+\beta_1*\mathscr{K}(\beta_0)+\beta_1*\mathscr{K}(\beta_0) \\ &-\beta_1*\mathscr{K}(\beta_0)*\mathscr{K}(\beta_0)\,.\end{aligned} \tag{A.11}$$

Proof. Using Lemma 6. □

Appendix B. Perturbed target system with approximate estimated kernel

We derive the perturbed target system (53), (54), where $w=u-\hat{k}*u$ and $\hat{k}$ is the approximate estimated kernel, assumed to be both continuous and differentiable with respect to t. Since (54) is just a consequence of the choice of $U(t)$, we focus on proving (53). Taking the derivative with respect to x and t of (52) gives the following

$$\begin{aligned}w_t &= u_t-\hat{k}_t*u-\hat{k}*u_t \\ w_x(x,t) &= u_x(x,t)-\hat{k}(0,t)u(x,t) \\ &\quad + \int_0^x \hat{k}_y(x-y,t)u(y,t)dy,\end{aligned} \tag{B.1}$$

for all $(x,t)\in[0,1]\times\mathbb{R}^+$. Integration by parts on (B.2) gives

$$w_x(x,t) = u_x(x,t)-\hat{k}(x,t)w(0,t)-\hat{k}*u_x(x,t)\,, \tag{B.2}$$

employing $u(0,t)=w(0,t)$. With (B.1), (2) and (B.2) gives

$$\begin{aligned}w_t(x,t)-w_x(x,t) = {}& w(0,t)\Big[\beta(x)+\hat{k}(x,t) \\ &-\int_0^x \hat{k}(x-y,t)\beta(y)dy\Big] \\ &-\hat{k}_t*u(x,t), \quad \forall(x,t)\in[0,1)\times\mathbb{R}^+.\end{aligned} \tag{B.3}$$

From (5), we have that

$$\hat{k} = -\hat{\beta}+\hat{\beta}*\hat{k}+\delta \tag{B.4}$$

and, with some rearrangements, arrive at

$$\begin{aligned}\beta(x)+\hat{k}(x,t)-\int_0^x \hat{k}(x-y,t)\beta(y)dy = {}&\tilde{\beta}(x,t)-\tilde{\beta}*\hat{k}(x,t) \\ &+\delta(x,t)\,.\end{aligned} \tag{B.5}$$

Then, using the inverse backstepping transformation $u=w-\hat{l}*w$, from (B.3) we arrive at (53).

Appendix C. Perturbed observer target system with exact estimated kernel

We derive the system (128)–(129). Since (129) is just a matter of the choice for the controller, we focus on (128). Taking the derivative of (127) with respect to t gives

$$w_t(x,t) = \hat{u}_t(x,t) - \int_0^x \breve{k}(x-y,t)\hat{u}_t(y,t)dy - \Omega(x,t)\,, \tag{C.1}$$

$$\Omega(x,t) := \int_0^x \breve{k}_t(x-y,t)\hat{u}(y,t)dy\,. \tag{C.2}$$

and with respect to x gives

$$\begin{aligned} w_x(x,t) &= \hat{u}_x(x,t) - \breve{k}(0,t)\hat{u}(x,t) + \int_0^x \breve{k}_y(y-x,t)\hat{u}(y,t)dy \\ &= \hat{u}_x(x,t) - \breve{k}(x,t)\hat{u}(0,t) - \int_0^x \breve{k}(x-y,t)\hat{u}_x(y,t)dy \end{aligned} \tag{C.3}$$

where we used integration by parts. Then gathering (C.1), (C.3) we have

$$\begin{aligned} w_t(x,t) - w_x(x,t) = {}& \hat{u}(0,t)\breve{k}(x,t) \\ & + u(0,t)\left[\hat{\beta}(x,t) - \int_0^x \breve{k}(x-y,t)\hat{\beta}(y,t)dy\right] \\ & - \Omega(x,t) + \gamma_0 u^2(0,t)\mathscr{B}(e,\breve{k})(x,t)\,. \end{aligned} \tag{C.4}$$

Using the definition of $\breve{k}$ in (5), as well as the inverse backstepping transformation, $\hat{u} = w - \hat{\beta} * w$, we arrive at (128).

References

Anfinsen, H., & Aamo, O. M. (2017). Adaptive output-feedback stabilization of linear 2 × 2 hyperbolic systems using anti-collocated sensing and control. *Systems & Control Letters*, *104*, 86–94.

Anfinsen, Henrik, & Aamo, Ole Morten (2018). Adaptive control of linear 2 × 2 hyperbolic systems. *Automatica*, *87*, 69–82.

Anfinsen, H., & Aamo, O. M. (2019). *Adaptive control of hyperbolic PDEs*. Springer.

Anfinsen, H., Diagne, M., Aamo, O. M., & Krstic, M. (2016). An adaptive observer design for $n+1$ coupled linear hyperbolic PDEs based on swapping. *IEEE Transactions on Automatic Control*, *61*, 3979–3990.

Bernard, Pauline, & Krstic, Miroslav (2014). Adaptive output-feedback stabilization of non-local hyperbolic PDEs. *Automatica*, *50*(10), 2692–2699.

Bhan, Luke, Bian, Yuexin, Krstic, Miroslav, & Shi, Yuanyuan (2024). PDE control gym: A benchmark for data-driven boundary control of partial differential equations. In Alessandro Abate, Kostas Margellos, & Antonis Papachristodoulou (Eds.), *Proceedings of machine learning research*: *vol. 242*, *Proceedings of the 6th annual learning for dynamics & control conference* (pp. 1–13). PMLR.

Bhan, Luke, Shi, Yuanyuan, & Krstic, Miroslav (2023). Neural operators for bypassing gain and control computations in PDE backstepping. *IEEE Transactions on Automatic Control*, 1–16, https://ieeexplore.ieee.org/document/10374221.

Bresch-Pietri, Delphine, & Krstic, Miroslav (2009). Adaptive trajectory tracking despite unknown input delay and plant parameters. *Automatica*, *45*(9), 2074–2081.

Bresch-Pietri, Delphine, & Krstic, Miroslav (2014). Output-feedback adaptive control of a wave PDE with boundary anti-damping. *Automatica*, *50*(5), 1407–1415.

Deng, Beichuan, Shin, Yeonjong, Lu, Lu, Zhang, Zhongqiang, & Karniadakis, George Em (2022). Approximation rates of DeepONets for learning operators arising from advection–diffusion equations. *Neural Networks*, *153*, 411–426.

Karafyllis, Iasson, & Krstic, Miroslav (2020). Stability results for the continuity equation. *Systems & Control Letters*, *135*, Article 104594.

Karafyllis, Iasson, Krstic, Miroslav, & Chrysafi, Katerina (2019). Adaptive boundary control of constant-parameter reaction–diffusion PDEs using regulation-triggered finite-time identification. *Automatica*, *103*, 166–179.

Krstic, Miroslav, Bhan, Luke, & Shi, Yuanyuan (2023). Neural operators of backstepping controller and observer gain functions for reaction-diffusion PDEs. arXiv preprint arXiv:2303.10506, https://arxiv.org/abs/2303.10506.

Krstic, M., Kanellakopoulos, I., & Kokotovic, P. (1995). *Nonlinear and adaptive control design*. John Wiley & Sons.

Krstic, Miroslav, & Smyshlyaev, Andrey (2008). Adaptive boundary control for unstable parabolic PDEs—Part I: Lyapunov design. *IEEE Transactions on Automatic Control*, *53*(7), 1575–1591.

Lamarque, Maxence, Bhan, Luke, Vazquez, Rafael, & Krstic, Miroslav (2024). Gain scheduling with a neural operator for a transport PDE with nonlinear recirculation. arXiv preprint arXiv:2401.02511, https://arxiv.org/abs/2401.02511.

Lanthaler, Samuel, Li, Zongyi, & Stuart, Andrew M. (2023). The nonlocal neural operator: Universal approximation. arXiv preprint arXiv:2304.13221, https://arxiv.org/abs/2304.13221.

Lanthaler, Samuel, Mishra, Siddhartha, & Karniadakis, George E. (2022). Error estimates for DeepONets: a deep learning framework in infinite dimensions. *Transactions of Mathematics and its Applications*, *6*(1), tnac001.

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

Lu, Lu, Meng, Xuhui, Mao, Zhiping, & Karniadakis, George Em (2021). DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, *63*(1), 208–228.

Mukherjee, Anirbit, & Roy, Amartya (2024). Size lowerbounds for deep operator networks. *Transactions on Machine Learning Research*.

Qi, Jie, Zhang, Jing, & Krstic, Miroslav (2023). Neural operators for delay-compensating control of hyperbolic PIDEs. arXiv preprint arXiv:2307.11436, https://arxiv.org/abs/2307.11436.

Smyshlyaev, Andrey, & Krstic, Miroslav (2007a). Adaptive boundary control for unstable parabolic PDEs—Part II: Estimation-based designs. *Automatica*, *43*(9), 1543–1556.

Smyshlyaev, Andrey, & Krstic, Miroslav (2007b). Adaptive boundary control for unstable parabolic PDEs—Part III: Output feedback examples with swapping identifiers. *Automatica*, *43*(9), 1557–1564.

Smyshlyaev, A., & Krstic, M. (2010). *Adaptive control of parabolic PDEs*. Princeton University Press.

Wang, Shanshan, Diagne, Mamadou, & Krstić, Miroslav (2023a). Deep learning of delay-compensated backstepping for reaction-diffusion PDEs. arXiv preprint arXiv:208.10501, https://arxiv.org/abs/2308.10501.

Wang, Shanshan, Diagne, Mamadou, & Krstić, Miroslav (2023b). Neural operator approximations of backstepping kernels for 2 × 2 hyperbolic PDEs. arXiv preprint arXiv:2312.16762, https://arxiv.org/abs/2312.16762.

Wang, Shanshan, Diagne, Mamadou, & Qi, Jie (2022). Delay-adaptive predictor feedback control of reaction–advection–diffusion PDEs with a delayed distributed input. *IEEE Transactions on Automatic Control*, *67*(7), 3762–3769.

Wang, Ji, & Krstic, Miroslav (2021). Adaptive event-triggered PDE control for load-moving cable systems. *Automatica*, *129*, Article 109637.

Wang, Ji, & Krstic, Miroslav (2023). Event-triggered adaptive control of coupled hyperbolic PDEs with piecewise-constant inputs and identification. *IEEE Transactions on Automatic Control*, *68*(3), 1568–1583.

Wang, Ji, Tang, Shu-Xia, & Krstic, Miroslav (2020). Adaptive output-feedback control of torsional vibration in off-shore rotary oil drilling systems. *Automatica*, *111*, Article 108640.

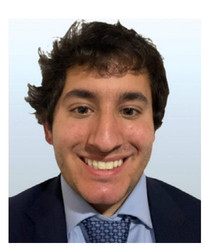
Wang, Ji, Tang, Shu-Xia, & Krstic, Miroslav (2023). Adaptive control of hyperbolic PDEs coupled with a disturbed and highly uncertain ODE. *IEEE Transactions on Automatic Control*, *68*(1), 108–123.

Yu, Huan, Vazquez, Rafael, & Krstic, Miroslav (2017). Adaptive output feedback for hyperbolic PDE pairs with non-local coupling. In *2017 American control conference* (pp. 487–492).

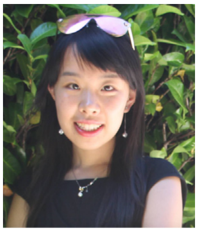
Zhang, Yihuai, Zhong, Ruiguo, & Yu, Huan (2023). Neural operators for boundary stabilization of stop-and-go traffic.

Zhu, Yang, & Krstic, Miroslav (2020). *Delay-adaptive linear control*. Princeton: Princeton University Press.

Maxence Lamarque is a graduate student at Ecole des Mines de Paris, France, due to graduate with a Master's degree in general engineering in 2025. He was a visiting student at University of California, San Diego, in Fall 2023.



Luke Bhan received his B.S. and M.S. degrees in Computer Science and Physics from Vander- bilt University in 2022. He is currently pursuing his Ph.D. degree in Electrical and Computer Engineering at the University of California, San Diego. His research interests include neural operators, learning-based control, and control of partial differential equations.



Yuanyuan Shi is an Assistant Professor of Electrical and Computer Engineering at the University of California, San Diego. She received her Ph.D. in Electrical Engineering, masters in Electrical Engineering and Statistics, all from the University of Washington, in 2020. From 2020 to 2021, she was a postdoctoral scholar at the California Institute of Technology. Her research interests include machine learning, dynamical systems, and control, with applications to sustainable power and energy systems.

Miroslav Krstic is Sr. Assoc. Vice Chancellor for Research at UC San Diego. He is Fellow of IEEE, IFAC, ASME, SIAM, AAAS, and Serbian Academy of Sciences and Arts. He has received the Bode Lecture Prize, Bellman Award, Reid Prize, Oldenburger Medal, Nyquist Lecture Prize, Paynter Award, Ragazzini Award, IFAC Ruth Curtain Distributed Parameter Systems Award, IFAC Nonlinear Control Systems Award, IFAC Adaptive and Learning System Award, Chestnut textbook prize, AV Balakrishnan Award, CSS Distinguished Member Award, the PECASE, NSF Career, and ONR YIP, Schuck ('96 and '19) and Axelby paper prizes.