Gain Scheduling with a Neural Operator for a Transport PDE with Nonlinear Recirculation

Maxence Lamarque, Luke Bhan, Rafael Vazquez, and Miroslav Krstic

***Abstract*— To stabilize PDE models, control laws require space-dependent functional gains mapped by nonlinear operators from the PDE functional coefficients. When a PDE is nonlinear and its "pseudo-coefficient" functions are state-dependent, a gain-scheduling (GS) nonlinear design is the simplest approach to the design of nonlinear feedback. The GS version of PDE backstepping employs gains obtained by solving a PDE at each value of the state. Performing such PDE computations in real time may be prohibitive. The recently introduced neural operators (NO) can be trained to produce the gain functions, rapidly in real time, for each state value, without requiring a PDE solution. In this paper we introduce NOs for GS-PDE backstepping. GS controllers act on the premise that the state change is slow and, as a result, guarantee only local stability, even for ODEs. We establish local stabilization of hyperbolic PDEs with nonlinear recirculation using both a "full-kernel" approach and the "gain-only" approach to gain operator approximation. Numerical simulations illustrate stabilization and demonstrate speedup by three orders of magnitude over traditional PDE gain-scheduling. Code (Github) for the numerical implementation is published to enable exploration.**

This work was the subject of the 2023 Bode Prize Lecture by the last coauthor.

I. INTRODUCTION

a) Goals and results: We present a model-based, gain-scheduled (GS) backstepping design for the control of a hyperbolic PDE with nonlinear recirculation, where the gain kernels are computed via a Neural Operator (NO). We leverage the recent breakthrough on deep neural network-based operator approximations (DeepONet [18]) to prove that our model-based GS design is stabilizable and real-time implementable due to an $\sim 10^3\times$ speedup thanks to the NO.

This is the first result in which the NO is used to recompute the kernel *online, at every time step*, based on learning a neural operator of the kernel only *once, offline*.

Even an exact gain scheduler is only locally stabilizing, and even local stabilization is complex to prove for a nonlinear PDE. We prove that both our exact and our NO-approximated GS-PDE backstepping controllers are stabilizing.

Maxence Lamarque is with École des Mines de Paris, France, maxence.lamarque@etu.minesparis.psl.eu

Luke, Bhan and Miroslav Krstic are with the University of California, San Diego, lbhan@ucsd.edu, krstic@ucsd.edu

Rafael Vazquez is with the Department of Aerospace Engineering Universidad de Sevilla, rvazquez1@us.es

The second author is supported by DOE grant DE-SC0024386 and the work of the fourth author was funded by FOSR grant FA9550-23-1-0535 and NSF grant ECCS-2151525.

b) Gain scheduling: In the GS literature for nonlinear ODEs we single out the classic papers [23], [24] and the survey [22], which makes links with linear parameter varying systems and highlights the need to limit the state derivative, with only local stabilization being generally possible with GS.

c) Stabilization of nonlinear PDEs using backstepping-based gain-scheduling: For the linear version of the hyperbolic PDEs we study here was introduced in [12]. The first and only attempt at global stabilization of a class of *general* nonlinear PDE systems was made in [26], [27], using infinite Volterra series. The complexity of that result highlights the importance of seeking methodologies that fall in between such complex nonlinear designs and linear designs such as [6]. In fact, for parabolic PDEs with nonlinear growth exceeding $|u|\log^2(1+|u|)$, [8] shows that finite-time blow up cannot be prevented with boundary control, hence, the pursuit of global stabilization is in vain, in general. A few CLF-based control designs for nonlinear PDEs include [10], [11]. Only a few particular cases admit a linearizing transformation [3], [15]. Hardly any results exist on GS for PDEs. The paper [2] focuses on finite-dimensional approximations. In [9], GS is used for time-varying (not nonlinear) PDEs. Our paper builds on [25], where exact—and explicit—GS for a special example of hyperbolic PDEs with recirculation is considered.

d) Learning approximations for model-based PDE control with stability guarantees: Recent breakthroughs in the mathematical properties of neural networks [5], [18], [16] have spawned an exploration of functional operator approximations in control theory. This was first explored for control of transport PDEs in [4] and extended to reaction-diffusion PDEs and observers in [14]. PDE backstepping [13], [12] kernels—typically governed by PDEs—were replaced with neural operator approximations. Under the universal approximation theorem [7], the papers [4], [14] prove that the PDE plant is stabilized provided the operator approximates the true kernel with required accuracy. Recently, [20] and [28] extended these initial results to hyperbolic and parabolic PDEs with delays, where the kernels are governed by multiple PDEs.

e) Paper outline: In Sec. II, we introduce the exact GS controller for the transport PDE with recirculation. In Sec. III, we establish the existence and continuours differentiability of the gain-schedulable backstepping kernel to be approximated by the DeepONet. In Sec. IV, we define the kernel operator and present its DeepONet approximation. In Section V, we present the main result: GS with DeepONet-approximated kernel is locally stabilizing. In Sec. VI and VII, we prove local stabilization in H_1 for the perturbed target system under the

DeepONet approximated kernel and then convert the perturbed system back into the original plant variables. In Sec. VIII, we present an alternative analysis that only employs only the evaluation of the gain kernel at the boundary and show the result is still locally stabilizing. Lastly, Sec. IX presents simulations illustrating the theoretical stabilization results.

f) Contributions: While [4] only gave a glimpse at possibilities of NOs, only linear PDEs have been considered so far. This paper takes up the NOs and GS simultaneously for the first time. Even our *exact GS* result, without the NO (obtained by setting $\epsilon = 0$ in our results and not belabored as a separate statement), is the first GS result for a nonlinear PDE, not counting the simple explicit GS example in [25]. When we combine GS with NO, two perturbations arise simultaneously in our analysis: (1) the perturbation from treating the state as quasi-constant, which even without an NO limits the achievable stabilization to local, and (2) the perturbation from the approximation of the GS kernel using an NO. Seeing the complexity of the analysis, the reader will understand why we pursue in this initial GS paper only a system where the "scheduling variable" is scalar (the state value at the outlet). Jumping straight into a general PDE class [26], [27], with a "scheduling variable" that is a *function* of space, would obscure the concept of simultaneously dealing with the PDE nonlinearities, GS design, and NO approximation. It would be incomprehensible and unpedagogical.

Our results include a DeepONet approximation theorem for a class of parametrized Volterra integral equations, H_1 local Lyapunov stability analyses for two approaches to approximation and backstepping transformation (the "full-kernel" approach, which employs the NO GS kernel in the backstepping transform, and the "gain-only" approach, which employs the exact GS kernel in backstepping), and simulations with a computational speedup of $\sim 10^3\times$ and stabilization.

g) Notation: We denote the convolution as $a * b(x,\nu) = \int_0^x a(x-y,\nu)b(y,\nu)dy$, and use the notation $a_{x\nu} = \frac{\partial^2 a}{\partial x \partial \nu}$ and $\|a(t)\| = \sqrt{\int_0^1 a^2(x,t)dx}$ for a function a on $[0,1]\times\mathbb{R}^+$.

II. Exact Gain-Scheduled (GS) PDE Backstepping for a Hyperbolic PDE with Nonlinear Recirculation

We study nonlinear hyperbolic PDE systems of the form

$$\begin{aligned} u_t(x,t) &= u_x(x,t) + \beta(x,u(0,t))u(0,t)\,, && (1)\\ u(1,t) &= U(t), \qquad (x,t)\in[0,1]\times\mathbb{R}^+\,, && (2)\end{aligned}$$

with the function $\beta \in \mathcal{C}^1([0,1]\times\mathbb{R},\mathbb{R})$ denoting the nonlinear, $u(0,t)$-dependent, recirculation gain. Our objective is to (locally) stabilize this system, at the equilibrium $u\equiv 0$, with a GS feedback implemented through the boundary input U. We "pretend," at each time t, that the state trace $u(0,t)$ is constant and employ a gain corresponding to such a "constant $u(0,t)$."

To obtain such a gain , wemap the system (1), (2) into a transport PDE $w_t = w_x$ using the backstepping transformation

$$w(x,t) = u(x,t) - \int_0^x k(x-y,u(0,t))u(y,t)dy\,,\qquad(3)$$

with a homogeneous inlet boundary condition $w(1,t)=0$, using the boundary control law

$$U(t) = \int_0^1 k(1-y,u(0,t))u(y,t)dy\,. \qquad (4)$$

If $u(0,t)$ were indeed a constant $\nu \equiv u(0,t)$, the kernel would need to satisfy the Volterra integral equation

$$k(x,\nu) = -\beta(x,\nu) + \int_0^x \beta(x-y,\nu)k(y,\nu)dy\,, \quad \forall (x,\nu)\in[0,1]\times\mathbb{R}\,. \qquad (5)$$

However, $u(0,t)$ is not constant and (1), (2) is not mapped into $w_t = w_x, w(1,t)=0$ but into a complex nonlinear PIDE

$$\begin{aligned} w_t(x,t) &= w_x(x,t) - w_x(0,t)\Omega(x,t), && (6)\\ w(1,t) &= 0, && (7)\end{aligned}$$

where

$$\begin{aligned} \Omega(x,t) &= \int_0^x [k_\nu(x-y,w(0,t))]\Big[w(y,t) \\ &\quad + \int_0^y l(y-s,w(0,t))w(s,t)ds\Big]dy && (8)\\ l(x,\nu) &= \int_0^x k(x-y,\nu)l(y,\nu)dy \\ &\quad + k(x,\nu), \quad (x,\nu)\in[0,1]\times\mathbb{R}\,. && (9)\end{aligned}$$

The system (6), (7), (8), (9) is called the "target system." Since the linearization of this nonlinear PIDE is the transport PDE $w_t = w_x, w(1,t)=0$, a lengthy Lyapunov analysis can be conducted which proves that the equilibrium of this target system is locally exponentially stable in the spatial H^1 norm.

In this paper, we apply an approximation of the GS controller (4) using deep learning. The result is a "perturbed target system," which is even more complex than (6), (7), (8), (9), and requires an even more complicated Lyapunov analysis.

III. Existence and Regularity of the "Gain-Schedulable" Backstepping Kernel $k(x,\nu)$

Lemma 1: [Existence and bound for kernel and its derivatives] For each $\beta\in\mathcal{C}^1([0,1]\times\mathbb{R})$, the Volterra equation (5),

$$k(x,\nu) = -\beta(x,\nu) + \int_0^x \beta(x-y,\nu)k(y,\nu)dy \qquad (10)$$

has a unique $\mathcal{C}^1([0,1]\times\mathbb{R})$ solution. In addition, when β is only defined for $(x,\nu)\in[0,1]\times[-B_\nu,B_\nu]$, $B_\nu>0$, the following holds on the same domain,

$$\begin{aligned} |k(x,\nu)| &\le B_\beta e^{B_\beta x}\,, && (11)\\ |k_\nu(x,\nu)| &\le \alpha_\nu e^{\alpha_\nu x}(1+\alpha_\nu x)\,, && (12)\end{aligned}$$

where $B_\beta := \|\beta\|_{\infty,[0,1]\times[-B_\nu,B_\nu]}$ and $\alpha_\nu := \max(\|\beta\|_{\infty,[0,1]\times[-B_\nu,B_\nu]}, \|\beta_\nu\|_{\infty,[0,1]\times[-B_\nu,B_\nu]})$. If, furthermore, $\beta_{x\nu}$ exists and is continuous, then $k_{x\nu}$ exists and is continuous. When defined on $[0,1]\times[-B_\nu,B_\nu]$, the second derivative satisfies, $\forall(x,\nu)\in[0,1]\times[-B_\nu,B_\nu]$,

$$|k_{x\nu}(x,\nu)| \le \alpha e^{\alpha x}(1+2\alpha) + \alpha^2 x e^{\alpha x}(\alpha+1), \qquad (13)$$

$$\alpha := \left\|\max\left(|\beta|,|\beta_\nu|,|\beta_{x\nu}|,|\beta_x|\right)\right\|_{\infty,[0,1]\times[-B_\nu,B_\nu]}\,. \qquad (14)$$

Proof: Existence and continuity of k: Let's recall the results proved in [4]. For each $\beta \in \mathcal{C}^0([0,1],\mathbb{R})$, the equation

$$k(x) = -\beta(x) + \int_0^x \beta(x-y)k(y)dy, \qquad x \in [0,1], \quad (15)$$

has a unique $\mathcal{C}^0([0,1],\mathbb{R})$ solution, which satisfies

$$k = \sum_{n=0}^{\infty} \Delta k^n \quad (16)$$

and $|k(x)| \le \bar{\beta} e^{\bar{\beta} x}$ for $x \in [0,1]$, where

$$\Delta k^{n+1}(x) := \int_0^x \beta(x-y)\Delta k^n(y)dy, \quad n \ge 0 \quad (17)$$

with $\Delta k^0 := -\beta$ and $\bar{\beta} := \|\beta\|_{\infty,[0,1]}$, and Δk^n satisfies

$$|\Delta k^n(x)| \le \frac{\bar{\beta}^{n+1} x^n}{n!}. \quad (18)$$

For $\nu \in \mathbb{R}$, since $\beta(\cdot,\nu)$ is continuous the previous statements ensure that there exists a continuous kernel $k(\cdot,\nu)$ that satisfies (15). Introducing the sequence

$$\begin{aligned} \Delta k^0 &:= -\beta, & (19)\\ \Delta k^{n+1} &:= \beta * \Delta k^n, \qquad x \in [0,1], & (20)\end{aligned}$$

it follows from (16) that for a given $\nu \in \mathbb{R}$,

$$\begin{aligned} k(x,\nu) &= \sum_{n=0}^{\infty} \Delta k^n(x,\nu), \qquad x \in [0,1], & (21)\\ |\Delta k^n(x,\nu)| &\le \frac{\bar{\beta}_\nu^{\,n+1} x^n}{n!}, \qquad x \in [0,1], & (22)\end{aligned}$$

where $\bar{\beta}_\nu := \|\beta(\cdot,\nu)\|_\infty$. Also, since β is continuous, (21), (22) give the continuity of k. With (22) and (21) we get (11).

Now let's prove that k is $\mathcal{C}^1$ by proving that both its partial derivatives exist and are continuous. To do so we prove that $\sum_{n=0}^{\infty} \Delta k_x^n, \sum_{n=0}^{\infty} \Delta k_\nu^n$ converge uniformly on $[0,1]\times[-B_\nu, B_\nu]$ for all $B_\nu > 0$.

Existence and continuity of k_x: For $n \in \mathbb{N}$, differentiating (20) with respect to x, since β is $\mathcal{C}^1$, gives the existence and continuity of $\Delta k_x^n, n \ge 0$ (for Δk^0 we use (19)). The details leading to this conclusion entail the study of

$$\begin{aligned} \Delta k_x^{n+1}(x,\nu) &= \beta(0,\nu)\Delta k^n(x,\nu) + \\ &\quad \int_0^x \beta_x(x-y,\nu)\Delta k^n(y,\nu)dy \quad (23)\end{aligned}$$

along with (22), which leads to the upper bound

$$|\Delta k_x^{n+1}(x,\nu)| \le \frac{\bar{\beta}_\nu^{n+2} x^n}{n!} + \alpha_x \frac{\bar{\beta}_\nu^{n+1} x^{n+1}}{(n+1)!}, \quad (24)$$

where $\alpha_x := \|\beta_x\|_{\infty,[0,1]\times[-B_\nu,B_\nu]} < \infty$ because β is assumed to be $\mathcal{C}^1$. The bound provided by (24) ensures that $\sum_{n=0}^{\infty} \Delta k_x^n$ converges uniformly on $[0,1]\times[-B_\nu,B_\nu]$. It follows that k_x exists and is continuous on $[0,1]\times\mathbb{R}$.

Existence and continuity of k_ν: We prove by induction that for all $n \in \mathbb{N}, \Delta k^n$ is differentiable with respect to ν, Δk_ν^n is continuous and that

$$|\Delta k_\nu^n(x,\nu)| \le \frac{(n+1)\alpha_\nu^{n+1} x^n}{n!}, \quad (25)$$

where $\alpha_\nu := \max\left(\|\beta\|_\infty, \|\beta_\nu\|_{\infty,[0,1]\times[-B_\nu,B_\nu]}\right) < \infty$. Since β is $\mathcal{C}^1$ we have that:

$$|\Delta k_\nu^0| \le \alpha_\nu. \quad (26)$$

Let $n \in \mathbb{N}$ such that the previous statement holds. Taking the derivative of (20) with respect to ν gives that, for $(x,\nu) \in [0,1]\times[-B_\nu,B_\nu]$,

$$\begin{aligned} \Delta k_\nu^{n+1}(x,\nu) &= \int_0^x (\beta_\nu(x-y,\nu)\Delta k^n(y,\nu) \\ &\quad + \beta(x-y,\nu)\Delta k_\nu^n(y,\nu))dy \quad (27)\end{aligned}$$

Using the upper bounds provided by (25) and (22) leads to the following inequality:

$$|\Delta k_\nu^{n+1}(x,\nu)| \le \frac{(n+2)\alpha_\nu^{n+2} x^{n+1}}{(n+1)!}. \quad (28)$$

The bound provided by (25) ensures that $\sum_{n=0}^{\infty} \Delta k_\nu^n$ converges uniformly on $[0,1]\times[-B_\nu,B_\nu]$, so (12) follows.

Existence and continuity of $k_{x\nu}$: If we also assume that $\beta_{x\nu}$ is defined and continuous, we prove that $k_{x\nu}$ exists, and is continuous, by proving that $\sum_{n=0}^{\infty} \Delta k_{x\nu}^n$ uniformly converges on $[0,1]\times[-B_\nu,B_\nu]$. Taking the derivative of (27) with respect to x gives that

$$\begin{aligned} \Delta k_{x\nu}^{n+1}(x,\nu) &= \beta_\nu(0,\nu)\Delta k^n(x,\nu) + \beta(0,\nu)\Delta k_\nu^n(x,\nu) \\ &\quad + \int_0^x (\beta_{x\nu}(x-y,\nu)\Delta k^n(y,\nu) \\ &\quad + \beta_x(x-y,\nu)\Delta k_\nu^n(y,\nu))dy. \quad (29)\end{aligned}$$

Introducing $\alpha := \max(\alpha_x, \alpha_\nu, \|\beta_{x\nu}\|_{\infty,[0,1]\times[-B_\nu,B_\nu]})$, using (25) and (22), (29) is bounded by

$$\begin{aligned} |\Delta k_{x\nu}^{n+1}(x,\nu)| &\le \frac{(n+2)\alpha^{n+2} x^n}{n!} + \frac{(n+2)\alpha^{n+2} x^{n+1}}{(n+1)!}, \\ &\qquad (x,\nu) \in [0,1]\times[-B_\nu,B_\nu]. \quad (30)\end{aligned}$$

The bound provided by (30) ensures that $\sum_{n=0}^{\infty} \Delta k_{x\nu}^n(x,\nu)$ converges uniformly on $[0,1]\times[-B_\nu,B_\nu]$ and leads us to the upper bound (13). ∎

IV. Neural Operators Approximate Gain-Scheduling Kernel k

Analytically solving (5) for arbitrary functions β, whose second argument is changing in real time, is impossible. We replace the exact solution k with a NO-approximated function $\hat{k}$. The question remains whether the control (4), using the approximation $\hat{k}$ instead of the exact kernel k, retains local exponential stability for the equilibrium $u \equiv 0$. We answer affirmatively as long as the approximation is 'close enough.' Towards that end, we need an approximation theorem.

Theorem 1 (DeepOnet universal approximation theorem [7]): 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 $\mathcal{U} : X \to U \subset \mathbb{R}^{d_u}$ and $\mathcal{V} : Y \to V \subset \mathbb{R}^{d_v}$ be sets of continuous functions $u(x)$ and $v(y)$, respectively. Let $\mathcal{U}$ also be compact. Assume the operator $\mathcal{G} : U \to V$ 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 $\theta^{(k)}$, $v^{(k)}$, neural networks $f_N(\cdot;\theta^{(k)})$, $g_N(\cdot;v^{(k)})$,

$k = 1, \ldots, p$, and $x_j \in X$, $j = 1, \ldots, m$, with corresponding $\mathbf{u}_m = (u(x_1), u(x_2), \ldots, u(x_m))^T$, such that

$$|\mathcal{G}(u)(y) - \mathcal{G}_{\mathbb{N}}(\mathbf{u}_m)(y)| < \epsilon \tag{31}$$

for all functions $u \in \mathcal{U}$ and all values $y \in Y$ of $\mathcal{G}(u)$. Where

$$\mathcal{G}_{\mathbb{N}}(y) = \sum_{k=1}^{p} g^{\mathcal{N}}(\mathbf{u}_m; v^{(k)}) f^{\mathcal{N}}(y; \theta^{(k)}) \tag{32}$$

The feedback gain to be scheduled, and approximated, is the output of an operator that we introduce next. To get to the output of that operator, we first need to introduce the set of (admissible) functions at the input of that operator. Based on Theorem 1, this set of functions needs to be compact. Various bounds on the operator's input functions therefore come into play. We introduce them next.

Let $B_\nu > 0$ and $\breve{B} = (B_\beta, B_{\beta_\nu}, B_{\beta_x}, B_{\beta_{x\nu}}) \in (\mathbb{R}_*^+)^4$. Let H denote the subset of $\mathcal{C}^1([0,1] \times [-B_\nu, B_\nu])$ such that, for all $\beta \in H$,

- $\beta, \beta_x, \beta_\nu, \beta_{x\nu}$ exist and are Lipschitz with the same Lipschitz constant.
- $\|\beta\|_\infty < B_\beta, \|\beta_\nu\|_\infty < B_{\beta_\nu}, \|\beta_x\|_\infty < B_{\beta_x}, \|\beta_{x\nu}\|_\infty < B_{\beta_{x\nu}}$

The reason behind assuming Lipschitzness of $\beta, \beta_x, \beta_\nu, \beta_{x\nu}$ is to ensure the compactness of H with the Arzelà-Ascoli theorem, which guarantees that the set of uniformly bounded and uniformly continuous functions is compact. Lipschitzness is a sufficient condition for uniform continuity and we require this simple property of β and its derivatives noted above.

We endow H with the norm

$$\|\beta\|_H := \|\beta\|_\infty + \|\beta_\nu\|_\infty + \|\beta_x\|_\infty + \|\beta_{x\nu}\|_\infty \,. \tag{33}$$

We denote by $\mathcal{K}$ the operator that maps β to the kernel k that satifies (5), i.e.,

$$k = \mathcal{K}(\beta) \,. \tag{34}$$

For the stability analysis explored later in the paper, developing a NO that approximates k alone is not enough. The perturbed target system obtained with a backstepping transformation with the approximated kernel has terms that contain the derivatives $\frac{\partial}{\partial x}, \frac{\partial}{\partial \nu}, \frac{\partial}{\partial \nu \partial x}$ of k. For this reason, the operator that we approximate with DeepONet must entail more than the operator $\mathcal{K}$. The more elaborate operator that we approximate is denoted by $\mathcal{M}$ and defined next.

Definition 1: The operator $\mathcal{M} : H \to \mathcal{C}^1([0,1] \times [-B_\nu, B_\nu], \mathbb{R}) \times \mathcal{C}^0([0,1] \times [-B_\nu, B_\nu], \mathbb{R})^3$ is defined by

$$\mathcal{M}(\beta) = (\mathcal{K}(\beta), \mathcal{K}_1(\beta), \mathcal{K}_2(\beta), \mathcal{K}_3(\beta)) \,, \tag{35}$$

and its last three components are defined as

$$\begin{aligned} \mathcal{K}_1(\beta) &= \frac{\partial}{\partial \nu}\mathcal{K}(\beta) \,, & (36)\\ \mathcal{K}_2(\beta) &= \frac{\partial}{\partial \nu}\frac{\partial}{\partial x}\mathcal{K}(\beta) \,, & (37)\\ \mathcal{K}_3(\beta) &= \frac{\partial}{\partial x}\mathcal{K}(\beta) \,. & (38) \end{aligned}$$

For a neural approximation of $\mathcal{K}$, Theorem 1 relies on the compactness of the input function space of $\mathcal{M}$, as well as on the operator's continuity. In the next lemma, we prove the Lipschitzness of $\mathcal{M}$, and thus its continuity.

Lemma 2: $\mathcal{M}$ is Lipschitz.

Proof: Lipschitzness of $\mathcal{K}$: It was shown in [4] that

$$\|k_1 - k_2\|_\infty \le C\|\beta_1 - \beta_2\|_\infty \,, \tag{39}$$

where $\beta_1, \beta_2 \in \mathcal{C}^0([0,1])$ such that $\|\beta_1\|_\infty, \|\beta_2\|_\infty < B, C := e^{3B}$ and k_1, k_2 are the solutions of (15). This leads to the Lipschitzness of $\mathcal{K}$ with e^{3B_β} as the Lipschitz constant.

$$\|\mathcal{K}(\beta_1) - \mathcal{K}(\beta_2)\|_\infty < C\|\beta_1 - \beta_2\|_\infty \,, \tag{40}$$

where $\beta_{1,2} \in H$.

Lipschitzness of $\mathcal{K}_1$: Let $\beta_1, \beta_2 \in H$ and $k_1 = \mathcal{K}(\beta_1)$, $k_2 = \mathcal{K}(\beta_2)$. From (5) it follows that:

$$\begin{aligned} \delta k_\nu &= -\delta\beta_\nu + \partial_\nu\beta_1 * \delta k + \delta\beta_\nu * k_2 \\ &\quad + \beta_1 * \delta k_\nu + \delta\beta * \partial_\nu k_2 & (41)\\ \delta k &= k_1 - k_2 \,, & (42)\\ \delta\beta &= \beta_1 - \beta_2 \,. & (43) \end{aligned}$$

Reproducing the successive approximation process, we introduce the sequence (δk_ν^n) that satisfies

$$\begin{aligned} \delta k_\nu^{n+1} &= \beta_1 * \delta k_\nu^n \,, & (44)\\ \delta k_\nu^0 &= -\delta\beta_\nu + \partial_\nu\beta_1 * \delta k + \delta\beta_\nu * k_2 \\ &\quad + \delta\beta * \partial_\nu k_2 \,. & (45) \end{aligned}$$

Using Lemma 1 and (40) gives the upper bounds

$$\begin{aligned} \|\delta k_\nu^0\|_\infty &\le A\|\delta\beta\|_H \,, & (46)\\ |\delta k_\nu^n(x,\nu)| &\le A\|\delta\beta\|_H \frac{B_\beta^n x^n}{n!} \,, & (47)\\ A(B_\beta, B_{\beta_\nu}) &> 0 \,. & (48) \end{aligned}$$

It follows that

$$\begin{aligned} \delta k_\nu &= \sum_{n=0}^{\infty} \delta k_\nu^n \,, & (49)\\ \|\delta k_\nu\|_\infty &\le A\|\delta\beta\|_H e^{B_\beta} \,. & (50) \end{aligned}$$

Lipschitzness of $\mathcal{K}_2$: Keeping the same notation as before, differentiating (5) with respect to ν and then with respect to x, we have that

$$\begin{aligned} \delta k_{x\nu}(x,\nu) &= -\delta\beta_{x\nu}(x,\nu) + \partial_\nu\beta_1(0,\nu)\delta k(x,\nu) \\ &\quad + \delta\beta_\nu(0,\nu)k_2(x,\nu) + \beta_1(0,\nu)\delta k_\nu(x,\nu) \\ &\quad + \delta\beta(0,\nu)\partial_\nu k_2(x,\nu) + \partial_\nu\partial_x\beta_1 * \delta k(x,\nu) \\ &\quad + \delta\beta_{x\nu} * k_2(x,\nu) + \partial_x\beta_1(x,\nu) * \delta k_\nu(x,\nu) \\ &\quad + \delta\beta_x * \partial_\nu k_2(x,\nu) \,, \\ &\qquad (x,\nu) \in [0,1] \times [-B_\nu, B_\nu] \,. & (51) \end{aligned}$$

Using Lemma 1, (40) and (50) gives the upper bound

$$\begin{aligned} \|\delta k_{x\nu}\|_\infty &\le B\|\delta\beta\|_H \,, & (52)\\ B(B_\beta, B_{\beta_\nu}, B_{\beta_x}, B_{\beta_{x\nu}}) &> 0 \,. & (53) \end{aligned}$$

Lipschitzness of $\mathcal{K}_3$: Differentiating (5) with respect to x leads to

$$\delta k_x(x,\nu) = \beta_1(0,\nu)\delta k(x,\nu) + \delta\beta(0,\nu)k_2(x,\nu)$$

$$
\begin{aligned}
&+\partial_x\beta_1 * \delta k(x,\nu) + \delta\beta_x * k_2(x,\nu) \\
&-\delta\beta_x(x,\nu)\,. \qquad (54)
\end{aligned}
$$

Using Lemma 1 and (40) gives the upper bound

$$
\begin{aligned}
\|\delta k_x\|_\infty &\leq D\|\delta\beta\|_H\,, && (55)\\
D(B_\beta, B_{\beta_x}) &> 0\,. && (56)
\end{aligned}
$$

The Lipschitzness of $\mathcal{K}, \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ gives the Lipschitzness of $\mathcal{M}$ (with, for example, a Lipschitz constant that is the maximum of the four Lipschitz constants). ∎

Using Lemma 2 and Theorem 1 we get the following.

Theorem 2: [Existence of a NO to approx. the kernel] For all $\beta \in H$ and $\epsilon > 0$, there exists a neural operator $\hat{\mathcal{K}}$ such that, for all $\forall(x,\nu) \in [0,1]\times[-B_\nu, B_\nu]$,

$$
\begin{aligned}
&|\mathcal{K}(\beta)(x,\nu) - \hat{\mathcal{K}}(\beta)(x,\nu)| \\
&+|\frac{\partial}{\partial\nu}(\mathcal{K}(\beta) - \hat{\mathcal{K}}(\beta))(x,\nu)| \\
&+|\frac{\partial}{\partial\nu}\frac{\partial}{\partial x}(\mathcal{K}(\beta) - \hat{\mathcal{K}}(\beta))(x,\nu)| \\
&+|\frac{\partial}{\partial x}(\mathcal{K}(\beta) - \hat{\mathcal{K}}(\beta))(x,\nu)| < \epsilon\,. \qquad (57)
\end{aligned}
$$

V. Main Result: Locally Stabilizing DeepONet Gain-Scheduling Feedback

Theorem 3: [Loc. stabilization by gain scheduling.] Let K, B_ν and the elements of the vector

$$\breve{B} = (B_\beta, B_{\beta_\nu}, B_{\beta_x}, B_{\beta_{x\nu}}), \qquad (58)$$

be positive and arbitrarily large. Then for all $c > 0$ there exist positive constants $\Omega_0(c,\breve{B},B_\nu) = \mathcal{O}_{c\to\infty}(ce^{-2c}), \epsilon^*(c,\breve{B},\Omega_0) = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}}), f(\breve{B}), M(c,\breve{B}) = f(\breve{B})e^c$ such that for any $\beta \in \mathcal{C}^1([0,1]\times\mathbb{R})$ with the properties that $\beta_{x\nu}$ is at least defined on $[0,1]\times[-B_\nu,B_\nu]$,

$$
\begin{aligned}
&|\beta(x,\nu)| \leq B_\beta, \quad |\beta_x(x,\nu)| \leq B_{\beta_x}, \quad |\beta_\nu(x,\nu)| \leq B_{\beta_\nu}, \\
&|\beta_{x\nu}(x,\nu)| \leq B_{\beta_{x\nu}}, \quad \forall(x,\nu)\in[0,1]\times[-B_\nu,B_\nu]\,, \quad (59)
\end{aligned}
$$

and that $\beta, \beta_x, \beta_\nu, \beta_{x\nu}$ are K-Lipschitz, any feedback law

$$U(t) = \int_0^1 \hat{k}(1-y, u(0,t))u(y,t)dy, \qquad (60)$$

with $\hat{k}$ being an approximated backstepping kernel provided by Theorem 2 for any accuracy $\epsilon \in (0,\epsilon^*)$, guarantees that, if the initial condition $u_0 := u(\cdot,0)$ of the system (1), (2) satisfies

$$\Omega(u_0) \leq \Omega_0\,, \qquad (61)$$

then

$$\Omega(u(t)) \leq M\Omega(u_0)e^{-\frac{c}{2}t}\,, \quad \forall t \geq 0\,, \qquad (62)$$

where

$$\Omega(u(t)) := u^2(0,t) + \|u(t)\|^2 + \|u_x(t)\|^2\,. \qquad (63)$$

This theorem is proven in Section VI, in which stability is studied for the perturbed target system, and Section VII, where the norm equivalence is established between the original and target system states, concluding at the end of Section VII.

VI. H_1 Lyapunov Analysis of Perturbed Target System

For $\beta \in H$, denote $\hat{k} := \hat{\mathcal{K}}(\beta)$ and let $\hat{\mathcal{K}}$ be any NO satisfying Theorem 2 for any accuracy $\epsilon > 0$. We consider now the "DeepONet-approximated" transform of (1),

$$w(x,t) = u(x,t) - \int_0^x \hat{k}(x-y, u(0,t))u(y,t)dy, \qquad (64)$$

which produces the following perturbed target system,

$$
\begin{aligned}
w_t(x,t) &= w_x(x,t) - w_x(0,t)\Omega(x,t) \\
&\quad +[\delta(x,w(0,t)) + \tilde{k}(0,w(0,t)\Omega(x,t))] \\
&\quad \times w(0,t), && (65)\\
w(1,t) &= 0, && (66)
\end{aligned}
$$

where

$$
\begin{aligned}
\Omega(x,t) &= \int_0^x [k_\nu(x-y,w(0,t)) + \delta_2(x-y,w(0,t))] \\
&\quad \times\Bigg[w(y,t) \\
&\quad + \int_0^y \hat{l}(y-s,w(0,t))w(s,t)ds\Bigg]dy\,, && (67)\\
\delta(x,\nu) &= -\tilde{k}(x,\nu) + \int_0^x \beta(x-y,\nu)\tilde{k}(y,\nu)dy\,, && (68)\\
\delta_2(x,\nu) &= -\tilde{k}_\nu(x,\nu)\,, && (69)\\
\hat{l}(x,\nu) &= \hat{k}(x,\nu) + \int_0^x \hat{k}(x-y,\nu)\hat{l}(y,\nu)dy\,, && (70)\\
\tilde{k} &= k - \hat{k}\,. && (71)
\end{aligned}
$$

Lemma 3: [Upper bound for inverse bkst kernel.] The inverse backstepping kernel $\hat{l}$ introduced as the solution of (70) satisfies the inequality

$$
\begin{aligned}
|\hat{l}(x,\nu)| &\leq (B_\beta + (1+B_\beta)\epsilon)e^{(1+B_\beta)\epsilon x}\,, && (72)\\
u(x,t) &= w(x,t) + \int_0^x \hat{l}(x-y,w(0,t))w(y,t)dy, && (73)
\end{aligned}
$$

$\forall(x,\nu)\in[0,1]\times[-B_\nu,B_\nu]$, where u is the state of plant (1), (2) and w is the state of the perturbed target system (65), (66).

Proof: First, notice that $\hat{k}$ satisfies

$$\hat{k} = -\beta + \beta * \hat{k} + \delta, \qquad (74)$$

where δ is defined in (68). Then, reusing the definition of $\hat{l}$ in (70), we have the following equality

$$\hat{l} = -\beta + \delta + \delta * \hat{l}\,. \qquad (75)$$

Applying, again, the successive approximation method, it holds that

$$
\begin{aligned}
|\hat{l}(x,\nu)| &\leq |B_\beta + \bar{\delta}|e^{\bar{\delta}x}\,, \\
\forall(x,\nu) &\in [0,1]\times[-B_\nu,B_\nu], && (76)
\end{aligned}
$$

where $\|\delta\|_\infty \leq \bar{\delta} := \epsilon(1+B_\beta)$. The proof of (73) is done by computing $w + \hat{l} * w$ with w satisfying (64), and using (70). ∎

Lemma 4: [Loc. Lyapunov estimate for perturbed target sys.] For any $\beta \in H$, for any $c > 0$ there exist positive functions $R_0(\breve{B},c,B_\nu) = \mathcal{O}_{c\to\infty}(ce^{-c})$, $\epsilon^*(\breve{B},c) = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}})$

with a decreasing dependence on argument $\breve{B}$ argument such that for all $\epsilon \in (0, \epsilon^*)$, the Lyapunov function

$$\begin{aligned} V(t) &= V_1(t) + V_2(t)\,, && (77)\\ V_1(t) &= \frac{1}{2}\int_0^1 e^{cx} w^2(x,t)dx\,, && (78)\\ V_2(t) &= \frac{1}{2}\int_0^1 e^{cx} w_x^2(x,t)dx\,, && (79)\end{aligned}$$

satifies :

$$V(0) \le R_0 \quad \Rightarrow \quad V(t) \le V(0)e^{-\frac{c}{2}t}, \qquad t \ge 0\,. \quad (80)$$

Proof: This proof is the central and longest part of the paper, establishing local exponential stability of the perturbed target system, robust to the specific neural network and the accuracy ϵ used to approximate the gain-scheduled control law.

Using Agmon's inequality on w satisfying (65), we have

$$|u(0,t)|^2 = |w(0,t)|^2 \le 2\|w(t)\|\,\|w_x(t)\| \le 2V(t)\,. \quad (81)$$

Then, if we chose the initial conditions small so that $V(0) \le \frac{B_\nu^2}{2}$ and prove that V is decreasing, the property $|u(0,t)| \le B_\nu$, important throughout our analysis, is enforced $\forall t \ge 0$.

Before starting the heavy computation of $\dot{V}$ we recall from Theorem 2 and Lemma 3 several upper bounds, and introduce additional ones, which that are used in the computation, which hold as a result of $|w(0,t)| \le B_\nu$ holding:

$$\begin{aligned} \|\delta(\cdot, w(0,t))\|_\infty &\le \epsilon(1+B_\beta) =: \bar{\delta}\,, && (82)\\ \|\delta_x(\cdot, w(0,t))\|_\infty &\le \epsilon(1+B_\beta+B_{\beta_x}) =: \bar{\delta_x}\,, && (83)\\ \|\delta_2(\cdot, w(0,t))\|_\infty &\le \epsilon\,, && (84)\\ \|\delta_{2x}(\cdot, w(0,t))\|_\infty &\le \epsilon\,, && (85)\\ \|\tilde{k}(\cdot, w(0,t))\|_\infty &\le \epsilon\,, && (86)\\ \|\hat{l}(\cdot, w(0,t))\|_\infty &\le (B_\beta+\bar{\delta})e^{\bar{\delta}} := \bar{l}\,, && (87)\\ \|k_\nu(\cdot, w(0,t))\|_\infty &\le \alpha e^{\alpha}(1+\alpha) =: \bar{k}_\nu\,, && (88)\\ \|k_{x\nu}(\cdot, w(0,t))\|_\infty &\le \alpha e^{\alpha}(1+2\alpha) \\ &\quad +\alpha^2 e^{\alpha}(\alpha+1) =: \bar{k}_{x\nu}, && (89)\end{aligned}$$

where

$$\alpha := \max(B_\beta, B_{\beta_x}, B_{\beta_\nu}, B_{\beta_{x\nu}})\,. \quad (90)$$

With these inequalities we have that

$$\begin{aligned} \|\Omega(\cdot,t)\|_\infty &\le \bar{\Omega}\int_0^1 |w(x,t)|dx, && (91)\\ |\Omega_x(x,t)| &\le \bar{\Omega}_{x1}\int_0^1 |w(y,t)|dy + \bar{\Omega}_{x2}|w(x,t)|, && (92)\end{aligned}$$

where

$$\begin{aligned} \bar{\Omega} &:= (\bar{k}_\nu+\epsilon)(\bar{l}+1)\,, && (93)\\ \bar{\Omega}_{x1} &:= \bar{l}(\bar{k}_\nu+\epsilon) + (1+\bar{l})(\bar{k}_{x\nu}+\epsilon)\,, && (94)\\ \bar{\Omega}_{x2} &:= \bar{k}_\nu+\epsilon\,. && (95)\end{aligned}$$

We prove (91) first. We assumed that $|w(0,t)| \le B_\nu$, then using triangular inequality, (88), (84), (76) we have that

$$\begin{aligned} |\Omega(x,t)| &\le (\bar{k}_\nu+\epsilon)\bigg(\int_0^x |w(y,t)|dy \\ &\quad +\bar{l}\int_0^x\int_0^y |w(s,t)|dsdy\bigg) \qquad (96)\end{aligned}$$

and just upper-bound it with $x, y = 1$. Turning to (92), we first take the derivative of (67) with respect to x,

$$\begin{aligned} \Omega_x(x,t) &= [k_\nu(0,w(0,t)) + \delta_2(0,w(0,t))](w(x,t) \\ &\quad + \int_0^x \hat{l}(x-s,w(0,t))w(s,t)ds) \\ &\quad + \int_0^x (k_{x\nu}(x-y,w(0,t)) \\ &\quad +\delta_{2x}(x-y,w(0,t))) \times \bigg[w(y,t) \\ &\quad + \int_0^y \hat{l}(y-s,w(0,t))w(s,t)ds\bigg]dy\,. \qquad (97)\end{aligned}$$

We arrive at (92) assuming that $|w(0,t)| \le B_\nu$ and using (88), (84), (76), (87) and (85). All of other upper bounds are obtained using Theorem 2 and Lemmas 1 and 3.

Estimate of V_1: Taking the derivative of (78), leads to the following equalities:

$$\dot{V}_1 = I_{11} + I_{12} + I_{13} + I_{14}, \quad (98)$$

where

$$\begin{aligned} I_{11}(t) &= -\frac{w^2(0,t)}{2} - \frac{c}{2}\int_0^1 e^{cx}w^2(x,t)dx\,, && (99)\\ I_{12}(t) &= -w_x(0,t)\int_0^1 e^{cx}w(x,t)\Omega(x,t)dx\,, && (100)\\ I_{13}(t) &= w(0,t)\int_0^1 e^{cx}w(x,t)\delta(x,t)dx\,, && (101)\\ I_{14}(t) &= w(0,t)\tilde{k}(0,w(0,t)) \\ &\quad \times\int_0^1 e^{cx}w(x,t)\Omega(x,t)dx\,. && (102)\end{aligned}$$

We first work on (100), $I_{12}(t)$. Using Young's inequality and (91) we arrive at the following inequality

$$I_{12}(t) \le \frac{w_x(0,t)^2}{8} + 4e^c\bar{\Omega}^2 V_1(t)\|w(t)\|^2\,. \quad (103)$$

Continuing with $I_{13}(t)$ in (101), using (82), Young's inequality and Cauchy-Scwarz inequalities we have the following upper bound

$$I_{13}(t) \le \frac{w(0,t)^2}{8} + 4\bar{\delta}^2 V_1(t)\frac{(e^c-1)\,.}{c} \quad (104)$$

Moving to $I_{14}(t)$ in (102), Young's inequality and (91) give the following inequality,

$$I_{14}(t) \le \frac{w(0,t)^2}{8} + 4\epsilon^2\bar{\Omega}^2 e^c\|w(t)\|^2 V_1(t)\,. \quad (105)$$

Gathering (99), (103), (104), (105) we arrive at the following inequality

$$\begin{aligned} \dot{V}_1(t) &\le \frac{w_x^2(0,t)}{8} - \frac{w^2(0,t)}{4} \\ &\quad -V_1(t)[c - 4\epsilon^2\bar{\Omega}^2 e^c\|w(t)\|^2 \\ &\quad -4\bar{\delta}^2\frac{e^c-1}{c} - 4e^c\bar{\Omega}^2\|w(t)\|^2]\,. \qquad (106)\end{aligned}$$

Estimate of V_2: Note that differentiating (64) with respect to x leads to the following PIDE system satisfied by w_x:

$$w_{xt}(x,t) = w_{xx}(x,t) - w_x(0,t)\Omega_x(x,t)$$

$$
\begin{aligned}
&\quad +\Big(\delta_x(x,w(0,t)) \\
&\quad +\tilde{k}(0,w(0,t)\Omega_x(x,t))\Big)w(0,t)\,, \qquad (107)\\
w_x(1,t) &= \Omega(1,t)w_x(0,t)-w(0,t)[\delta(1,w(0,t)) \\
&\quad +\tilde{k}(0,w(0,t))\Omega(1,t)]\,. \qquad (108)
\end{aligned}
$$

Taking the derivative of (79) leads to the following:

$$
\begin{aligned}
\dot{V}_2 &= I_{21}+I_{22}+I_{23}+I_{24}\,, && (109)\\
I_{21}(t) &= \frac{e^c}{2}w_x^2(1,t)-\frac{w_x^2(0,t)}{2} \\
&\quad -\frac{c}{2}\int_0^1 e^{cx}w_x^2(x,t)dx\,, && (110)\\
I_{22}(t) &= -w_x(0,t)\int_0^1 e^{cx}w_x(x,t)\Omega_x(x,t)dx\,, && (111)\\
I_{23}(t) &= w(0,t)\int_0^1 e^{cx}w_x(w,t)\delta_x(x,w(0,t))dx, && (112)\\
I_{24}(t) &= w(0,t)\tilde{k}(0,w(0,t)) \\
&\quad \times\int_0^1 e^{cx}w_x(w,t)\Omega_x(x,t)dx\,. && (113)
\end{aligned}
$$

Starting with $I_{21}(t)$. We'll specifically bound the term $w_x^2(1,t)$. But first let's notice that thanks to Hölder's inequality we have the upper bound

$$
\int_0^1 |w(x,t)|dx \le \|w(t)\| \le \sqrt{2V(t)}\,. \qquad (114)
$$

Using (108), (91), (82), (86) and (114) we have the following inequality for $w_x^2(1,t)$

$$
\begin{aligned}
w_x^2(1,t) &\le 2w_x^2(0,t)\bar{\Omega}^2V(t) \\
&\quad +w^2(0,t)\epsilon^2(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2+\epsilon^2w^2(0,t) \\
&\quad +2V(t)w_x^2(0,t)\bar{\Omega}^2(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2\,. \qquad (115)
\end{aligned}
$$

Then using (110), (115) gives

$$
\begin{aligned}
I_{21}(t) &\le \frac{e^c}{2}[2w_x^2(0,t)\bar{\Omega}^2V(t) \\
&\quad +w^2(0,t)\epsilon^2(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2+\epsilon^2w^2(0,t) \\
&\quad +2V(t)w_x^2(0,t)\bar{\Omega}^2(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2] \\
&\quad -\frac{w_x^2(0,t)}{2}-cV_2(t)\,. \qquad (116)
\end{aligned}
$$

Next, we move on to $I_{22}(t)$ in (111). Using (92) as well as Young's inequality and Cauchy-Schwarz inequality we have that

$$
I_{22}(t) \le \frac{w_x(0,t)^2}{4}+4e^cV_2(t)\|w(t)\|^2(\bar{\Omega}_{x1}^2+\bar{\Omega}_{x2}^2)\,. \qquad (117)
$$

Advancing to $I_{23}(t)$ in (112), using Young's inequality, (83) and Cauchy-Schwarz inequality we obtain

$$
I_{23}(t) \le \frac{w^2(0,t)}{8}+4V_2(t)\bar{\delta_x}^2\frac{(e^c-1)}{c}\,. \qquad (118)
$$

Finally for $I_{24}(t)$ in (113), using Young's inequality, (92), (86) and the Cauchy-Schwarz inequality leads to

$$
I_{24}(t) \le \epsilon^2\frac{w^2(0,t)}{4}+4e^cV_2(t)\|w(t)\|^2(\bar{\Omega}_{x1}^2+\bar{\Omega}_{x2}^2)\,. \qquad (119)
$$

Using the inequalities (116), (117), (118) and (119) gives

$$
\begin{aligned}
\dot{V}_2(t) &\le w^2(0,t)\Big[\frac{\epsilon^2}{4}+\frac{1}{8} \\
&\quad +\epsilon^2\frac{e^c}{2}(1+(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2)\Big] \\
&\quad -w_x^2(0,t)\Big[\frac{1}{4}-e^cV(t) \\
&\quad \times\bar{\Omega}^2(1+(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2))\Big] \\
&\quad -V_2(t)\Big[c-8e^c(\bar{\Omega}_{x1}^2+\bar{\Omega}_{x2}^2)\|w(t)\|^2 \\
&\quad -4\bar{\delta}_x^2\frac{(e^c-1)}{c}\Big]\,. \qquad (120)
\end{aligned}
$$

In summary, gathering (106) and (120) we have

$$
\begin{aligned}
\dot{V}(t) &\le -w(0,t)^2\Big[\frac{1}{8}-\frac{\epsilon^2}{4} \\
&\quad -\frac{\epsilon^2e^c}{2}(1+(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2)\Big] \\
&\quad -w_x(0,t)^2\Big[\frac{1}{8}-e^cV(t) \\
&\quad \times\bar{\Omega}^2(1+(1+B_\beta+\sqrt{2V(t)}\bar{\Omega})^2)\Big] \\
&\quad -V_1(t)\Big[c-4\epsilon^2\bar{\Omega}^2e^c\|w(t)\|^2 \\
&\quad -4\bar{\delta}^2\frac{e^c-1}{c}-4e^c\bar{\Omega}^2\|w(t)\|^2\Big] \\
&\quad -V_2(t)\Big[c-8e^c(\bar{\Omega}_{x1}^2+\bar{\Omega}_{x2}^2)\|w(t)\|^2 \\
&\quad -4\bar{\delta}_x^2\frac{e^c-1}{c}\Big]\,. \qquad (121)
\end{aligned}
$$

We kept using $\bar{\Omega}, \bar{\Omega}_{x1}, \bar{\Omega}_{x2}$ without specifying that they are functions of ϵ. In what follows we restrict, for simplicity, the choice of ϵ to $\epsilon\in(0,1]$ and use the notation $\bar{\Omega} := \bar{\Omega}(1), \bar{\Omega}_{x1} := \bar{\Omega}_{x1}(1), \bar{\Omega}_{x2} := \bar{\Omega}_{x2}(1)$. We emphasize that all the previous inequalities are valid for all $\epsilon\le 1$. We introduce the $\mathcal{K}_\infty$ function

$$
\beta_2(V) := e^c\bar{\Omega}^2V+e^cV\bar{\Omega}^2(1+B_\beta+\sqrt{2V}\bar{\Omega})^2\,, \qquad (122)
$$

and the quantity $R_1 := \beta_2^{-1}(1/8) > 0$. That way, if we assume that $V(t)\le R_1$, then (121) is negative. Moving on to (121), we introduce

$$
\epsilon_1 := \min\left(1, \frac{1}{\sqrt{2+4e^c(1+(1+B_\beta+\sqrt{2R_1}\bar{\Omega})^2)}}\right)\,. \qquad (123)
$$

If we take $\epsilon\le\epsilon_1$, and $V\le R_1$ we have that (121) is negative. Advancing to (121), we introduce the positive constants

$$
R_2 := \min\left(R_1, \frac{ce^{-c}}{32\bar{\Omega}^2}\right) > 0\,, \qquad (124)
$$

$$
\epsilon_2 := \min\left(\epsilon_1, \sqrt{\frac{\frac{c}{2}-8e^c\bar{\Omega}^2R_2}{8\bar{\Omega}^2e^cR_2+4(1+B_\beta)^2\times\frac{e^c-1}{c}}}\right)\,. \qquad (125)
$$

With these definitions, if we choose $\epsilon \leq \epsilon_2, V(0) \leq R_2$, we have (121), (121) negative and (121) is bounded by $-\frac{c}{2}V_1$. This is achieved by noticing that $\|w(t)\|^2 \leq 2V(t)$. Finally, moving on to (121), we introduce

$$R_3 := \min\left(R_2, \frac{ce^{-c}}{64(\bar{\Omega}_{x1}^2 + \bar{\Omega}_{x2}^2)}\right) > 0\,, \tag{126}$$

$$\epsilon^* := \min\Bigg(\epsilon_2, \frac{1}{2(1+B_\beta+B_{\beta_x})} \times\sqrt{\frac{\frac{c^2}{2} - 16ce^c(\bar{\Omega}_{x1}^2 + \bar{\Omega}_{x2}^2)R_3}{(e^c-1)}}\Bigg) > 0\,, \tag{127}$$

with which we get that, if we choose $\epsilon \leq \epsilon^*, V(0) \leq R_3$, we have that (121), (121) are negative, (121) is bounded by $-\frac{c}{2}V_1$, and (121) is bounded by $-\frac{c}{2}V_2$. To recapitulate our many majorizations, choosing $\epsilon \leq \epsilon^*$ and

$$V(0) \leq \min\left(R_3, \frac{B_\nu^2}{2}\right) =: R_0\,. \tag{128}$$

we arrive at the inequality

$$\dot{V}(t) \leq -\frac{c}{2}V, \qquad t \geq 0\,. \tag{129}$$

Using the comparison lemma we complete the proof. ∎

Lemma 5: [Loc. e.s. of perturbed target sys.] For all $c > 0$, there exist functions $\epsilon^*(\breve{B}, c) = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}})$, $\Psi_0(B_\nu, \breve{B}, c) = \mathcal{O}_{c\to\infty}(ce^{-2c}) > 0$ with a decreasing dependence in the argument $\breve{B}$ such that, for any $\beta \in H$ and any NO satisfying Theorem 2 for that β with any $\epsilon \in (0, \epsilon^*)$, where ϵ^* is defined in (127), if the perturbed target system (65), (66) is initialized with $w_0 := w(\cdot, 0)$ such that $\Psi(w_0) \leq \Psi_0$, then

$$\Psi(w(t)) \leq e^c\Psi(w_0)e^{-\frac{c}{2}t}\,, \qquad t \geq 0\,. \tag{130}$$

where

$$\Psi(w(t)) := \|w(t)\|^2 + \|w_x(t)\|^2\,. \tag{131}$$

Proof: Using the Lyapunov function defined in (77) for $c > 0$, Lemma 4 ensures the existence of $R_0, \epsilon^* > 0$ such that if $V(0) \leq R_0,\ \epsilon \leq \epsilon^*$

$$V(t) \leq V(0)e^{-\frac{c}{2}t}\,. \tag{132}$$

Note that

$$\Psi(w(t)) \leq 2V(t) \leq e^c\Psi(w(t))\,. \tag{133}$$

Choosing

$$\Psi_0 := 2e^{-c}R_0\,, \tag{134}$$

and ϵ^* as in (127) completes the proof. ∎

VII. Local Stability in Original Plant Variable

Lemma 6: [Equiv. norm perturb. target system] There exist $\rho, \delta > 0$ such that $\forall t \geq 0$, if $|u(0,t)| \leq B_\nu$ then,

$$\Psi(w(t)) \leq \delta\,\Omega(u(t))\,, \tag{135}$$

$$\Omega(u(t)) \leq \rho\,\Psi(w(t))\,. \tag{136}$$

Proof: We use Lemma 1, (11). We thus have that

$$|\hat{k}(x,\nu)| \leq \epsilon + B_\beta e^{B_\beta} =: \bar{k}\,, \tag{137}$$

$$|\hat{k}_x(x,\nu)| \leq \epsilon + (B_{\beta_x} + B_\beta^2 e^{B_\beta} + B_{\beta_x}B_\beta e^{B_\beta}) =: \bar{k}_x\,, \tag{138}$$

$\forall(x,\nu) \in [0,1]\times[-B_\nu, B_\nu]$, where $\hat{k} := \hat{\mathcal{K}}(\beta)$. With the assumption that $|u(0,t)| \leq B_\nu$, we apply the inequalities (137), (138) for $\hat{k}(x, u(0,t)), \hat{k}_x(x, u(0,t))$. Using (64), (137), Young's, and the Cauchy-Schwarz inequalities, we have that

$$\|w(t)\|^2 \leq 2\|u(t)\|^2 + 2\bar{k}^2\|u(t)\|^2\,. \tag{139}$$

Using (64), (138), (137), the Cauchy-Schwarz, and Young's inequality we have that

$$\|w_x(t)\|^2 \leq 4\|u_x(t)\|^2 + 4\bar{k}_x^2\|u(t)\|^2 + 2\bar{k}^2\|u(t)\|^2\,. \tag{140}$$

Gathering (139) and (140) we have that

$$\Psi(w(t)) \leq (6 + 4\bar{k}^2 + 4\bar{k}_x^2)\Omega(u(t)) =: \delta\,\Omega(u(t))\,. \tag{141}$$

Since $u(0,t) = w(0,t)$, we still have $|w(0,t)| \leq B_\nu$. Before constructing ρ, let's notice that taking the derivative of (70) with respect to x gives

$$\hat{l}_x(x, w(0,t)) = \hat{k}_x(x, w(0,t)) + \hat{k}(0, w(0,t))\hat{l}(x, w(0,t)) + \int_0^x \hat{k}_x(x-y, w(0,t))\hat{l}(y, w(0,t))dy\,. \tag{142}$$

Then using (137), (138), (76) we have that

$$|\hat{l}_x(x, w(0,t))| \leq \bar{k}_x + \bar{k}\bar{l} + \bar{k}_x\bar{l} =: \bar{l}_x\,. \tag{143}$$

Let's also notice that

$$u^2(0,t) = w^2(0,t) \leq \Psi(w(t))\,. \tag{144}$$

Using (144), (76), (143), and (73) it is shown the same way as before that

$$\Omega(u(t)) \leq (7 + 4\bar{l}^2 + 4\bar{l}_x^2)\Psi(w(t)) =: \rho\,\Psi(w(t))\,. \tag{145}$$

The constants ρ, δ increase in ϵ. Recalling (123), $\epsilon \leq \epsilon^* \leq 1$. To remove the dependency of δ, ρ on ϵ, we take their values at $\epsilon = 1$ and all the previous inequalities remain valid. ∎

With the lemmas, we complete the proof of Theorem 3.

Proof: [of Theorem 3] From Lemma 5 there exist $\epsilon^*(\breve{B}, c), \Psi_0(\breve{B}, B_\nu, c) > 0$, such that if $\Psi(w_0) \leq \Psi_0$, $\epsilon \in (0, \epsilon^*)$ then

$$\Psi(w(t)) \leq e^c\Psi(w_0)e^{-\frac{c}{2}t}, \qquad t \geq 0\,. \tag{146}$$

Recalling (81), (128), (133) and (134) If $\Psi(w_0) \leq \Psi_0$ then we also have that

$$u^2(0,t) \leq B_\nu^2, \qquad \forall t \geq 0\,. \tag{147}$$

Then we introduce the quantity $\Omega_0 := \min(\frac{\Psi_0}{\delta}, B_\nu^2)$, where δ is defined in Lemma 6. We choose $\Omega(u_0)$ such that $\Omega(u_0) \leq \Omega_0$. Since $u(0,0)^2 \leq \Omega_0 \leq B_\nu^2$, we use inequality (135), which ensures $\Psi(w_0) \leq \delta\Omega(u_0) \leq \Psi_0$. Since (147) is now valid, we use (136) which ensures that

$$\Omega(u(t)) \leq \rho\Psi(w(t)) \leq \rho e^c\Psi(w_0)e^{-c^*t} \leq \rho e^c\delta\Omega(u_0)e^{-c^*t}\,, \tag{148}$$

with ρ, δ defined in Lemma 6, completing the proof of local exponential stability in H^1 asserted in Theorem 3. ∎

VIII. A Gain-Only Approach to Approximate Gain Scheduling

In this section we present an alternative approach to the approximation of the kernel and the stability analysis. Instead of approximating the four functions $k, k_\nu, k_x, k_{x\nu}$, as per Definition 1 and Lemma 2, we approximate only the functions k, k_ν. This relaxes the DeepONet training requirements.

This training reduction comes at a price in analysis. Instead of the approximate transformation (64), which employs the kernel $\hat{k}$, we use the *exact* transformation (3) with kernel k, which, along with the same feedback law U defined in (4), maps the system (1), (2) into the nonlinear PIDE

$$w_t(x,t) = w_x(x,t) - w_x(0,t)\Omega(x,t), \tag{149}$$
$$w(1,t) = \Gamma(t), \tag{150}$$

where

$$\Gamma(t) = -\int_0^1 \tilde{k}(1-y, w(0,t))\Big[w(y,t) \tag{151}$$
$$+\int_0^y l(y-s, w(0,t))w(s,t)ds\Big]dy, \tag{152}$$

and for the reader's convenience we recall (8) and (9)

$$\Omega(x,t) = \int_0^x k_\nu(x-y, w(0,t))\Big[w(y,t) + \int_0^y l(y-s, w(0,t))w(s,t)ds\Big]dy, \tag{153}$$
$$l(x,\nu) = k(x, w(0,t)) + \int_0^x k(x-y, w(0,t))l(y, w(0,t))dy \tag{154}$$

A close comparison of (149) with (65), as well as (153) with (67), reveals that the pertubation of the target system in the PDE domain no longer contains a perturbation based on the kernel approximation but retains a perturbation due to the GS-induced nonlinearity, (149). Additionally, by comparing (150) with (66), we see that the boundary condition is now perturbed, by Γ defined in (152). The consequence of the approximation-based perturbation moving from the PDE domain into the boundary is that the difficulty of the Lypunov analysis increases somewhat. A change of the norm is needed because of the perturbation in the boundary condition (150). We cannot ensure through Agmon's inequality alone that $w(0,t)$ remains bounded since $w(1,t) \neq 0$ which was a simplifying aspect of the proof of Lemma 4. Instead of the norm $\|w(t)\|^2 + \|w_x(t)\|^2$, $w(t) \in \mathcal{C}^1([0,1])$, we work with

$$\Psi(w(t)) = w^2(0,t) + \|w(t)\|^2 + \|w_x(t)\|^2 . \tag{155}$$

A. Relaxed kernel approximation

We redefine the set H and the NO used to approximate k. We relax H to the subset of $\mathcal{C}^1([0,1] \times [-B_\nu, B_\nu])$ such that for all $\beta \in H$,

- $\beta, \beta_x, \beta_\nu, \beta_{x\nu}$ exist and β, β_x are Lipschitz with the same Lipschitz constant,
- $\|\beta\|_\infty < B_\beta, \|\beta_\nu\|_\infty < B_{\beta_\nu}, \|\beta_x\|_\infty < B_{\beta_x}, \|\beta_{x\nu}\|_\infty < B_{\beta_{x\nu}}$.

Note that with the new H, we aren't required to have $\beta_x, \beta_{x\nu}$ Lipschitz. The reason behind this choice is that we don't approximate $k_x, k_{x\nu}$ but only k, k_ν. We don't require $\beta_x, \beta_{x\nu}$ uniformly bounded for the approximation but so that feedback will be available to all functions $\beta \in H$. With this new set H we require a weaker version of Theorem 2.

Theorem 4: [Neural operator to only approximate k, k_ν.] Given the operator $\mathcal{K}$

$$\mathcal{K} : \beta \to k(\beta) , \tag{156}$$

for all $\beta \in H$ and $\epsilon > 0$, there exists a neural operator $\hat{\mathcal{K}}$ such that, $\forall (x,\nu) \in [0,1] \times [-B_\nu, B_\nu]$,

$$|\mathcal{K}(\beta)(x,\nu) - \hat{\mathcal{K}}(\beta)(x,\nu)| + |\frac{\partial}{\partial\nu}(\mathcal{K} - \hat{\mathcal{K}})(x,\nu)| < \epsilon . \tag{157}$$

Having introduced the relaxed operator, we proceed to achieving a stabilization result analogous to Theorem 3.

Theorem 5: [Loc. stabilization by gain scheduling.] Let K, B_ν and the elements of the vector

$$\breve{B} = (B_\beta, B_{\beta_{\nu x}}, B_{\beta_x}, B_{\beta_{x\nu}}) , \tag{158}$$

be positive and arbitrarily large. Then for all $c > 0$ there exist positive constants $\Omega_0(c, \breve{B}, B_\nu) = \mathcal{O}_{c\to\infty}(e^{-2c})$, $\epsilon^*(c, \breve{B}, \Omega_0) = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}}), f(\breve{B}), M(c, \breve{B}) = f(\breve{B})ce^c$ such that for any $\beta \in \mathcal{C}^1([0,1] \times \mathbb{R})$ with the properties that $\beta_{x\nu}$ is at least defined on $[0,1] \times [-B_\nu, B_\nu]$,

$$|\beta(x,\nu)| \leq B_\beta, \tag{159}$$
$$|\beta_x(x,\nu)| \leq B_{\beta_x}, \tag{160}$$
$$|\beta_\nu(x,\nu)| \leq B_{\beta_\nu}, \tag{161}$$
$$|\beta_{x\nu}(x,\nu)| \leq B_{\beta_{x\nu}}, \tag{162}$$
$$\forall (x,\nu) \in [0,1] \times [-B_\nu, B_\nu] ,$$

and that β, β_ν are K-Lipschitz, any feedback law

$$U(t) = \int_0^1 \hat{k}(1-y, u(0,t))u(y,t)dy, \tag{163}$$

with $\hat{k}$ being an approximated backstepping kernel provided by Theorem 4 for any accuracy $\epsilon \in (0, \epsilon^*)$, guarantees that, if the initial condition $u_0 := u(\cdot, 0)$ of system (1), (2) satisfies

$$\Omega(u_0) \leq \Omega_0 , \tag{164}$$

then

$$\Omega(u(t)) \leq M\Omega(u_0)e^{-\frac{c}{2}t} , \quad \forall t \geq 0 , \tag{165}$$

where

$$\Omega(u(t)) := u^2(0,t) + \|u(t)\|^2 + \|u_x(t)\|^2 . \tag{166}$$

Notice that we now have a quantitatively slightly weaker result than the one stated in Lemma 3: an overshooting coefficient proportional to ce^c instead of e^c, and, additionally, the restriction on Ψ_0 has changed from $\mathcal{O}_{c\to\infty}(ce^{-2c})$ to the more conservative $\mathcal{O}_{c\to\infty}(e^{-2c})$. This slight weakening of the result follows from not having to approximate $k_{x\nu}, k_x$.

This theorem is proven in Section VIII-B, in which the stability is studied for the perturbed target system, and section VIII-C where the norm equivlence is established between the original and target system states.

B. Lyapunov analysis of target system with perturbed boundary conditions

Before Lyapunov analysis, we state bounds for the kernel of the inverse transformation, which is the counterpart of Lemma 3, but with the exact inverse backstepping kernel l instead.

Lemma 7: [Upper bound the for the exact inverse backstepping kernel and its derivative] The inverse backstepping transform's kernel l, defined in (9), satisfies the upper bounds

$$|l(x,\nu)| \leq \bar{k}e^{\bar{k}} =: \bar{l}\,, \tag{167}$$
$$|l_\nu(x,\nu)| \leq \bar{k}_\nu(1+\bar{l})e^{\bar{k}} =: \bar{l}_\nu\,, \tag{168}$$

for every $(x,\nu) \in [0,1]\times[-B_\nu, B_\nu]$ and where

$$\bar{k} := B_\beta e^{B_\beta}\,, \tag{169}$$

and $\bar{k}_\nu$ is defined in (88)

Proof: We use Lemma 1 and the successive approximation method to achieve (167) and the same for (168) noticing that taking the derivative of (9) with respect to ν gives

$$\begin{aligned} l_\nu(x,\nu) &= k_\nu(x,\nu) + \int_0^x \Big[k_\nu(x-y,\nu)l(y,\nu) \\ &\quad + k(x-y,\nu)l_\nu(y,\nu)\Big]dy\,, \end{aligned} \tag{170}$$

for all $(x,\nu)\in[0,1]\times[-B_\nu,B_\nu]$. ■

Lemma 8: [Loc. Lyapunov estimate pert. boundary conditions] For any $\beta \in H$, for any $c \geq 1$ there exist positive functions $R_0(\breve{B}, c, B_\nu) = \mathcal{O}_{c\to\infty}(ce^{-c}), \epsilon^* = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}})$ with a decreasing dependence on argument $\breve{B}$ such that for all $\epsilon \in (0,\epsilon^*)$ the Lyapunov function

$$V(t) = V_1(t) + V_2(t) + V_3(t)\,, \tag{171}$$
$$V_1(t) = \frac{c}{2}\int_0^1 e^{cx}w^2(x,t)dx\,, \tag{172}$$
$$V_2(t) = \frac{c}{2}\int_0^1 e^{cx}w_x^2(x,t)dx\,, \tag{173}$$
$$V_3(t) = \frac{1}{8}w^2(0,t)\,, \tag{174}$$

satisfies

$$V(0)\leq R_0 \implies V(t)\leq V(0)e^{-\frac{c}{2}t}, \quad t\geq 0\,. \tag{175}$$

The result also holds for $c\in(0,1)$ but we don't spend the time proving it since it is for large c (rapid decay) that the result is of interest. If it were interested to prove the result for $c\in(0,1)$, it would be easier to use the Lyapunov function

$$\begin{aligned} V(t) &= \frac{1}{2}\int_0^1 (1+x)w^2(x,t)dx \\ &\quad + \frac{1}{2}\int_0^1 (1+x)w_x^2(x,t)dx + \frac{w^2(0,t)}{8}\,. \end{aligned} \tag{176}$$

We would adapt the entire proof removing all terms in c and get a bound in the form $V(t) \leq V(0)e^{-\frac{t}{2}}$. Noticing that $e^{-\frac{t}{2}} \leq e^{\frac{-ct}{2}}$ we would have the result valid for all $c>0$.

With the definition used for $c\geq 1$, we have the equivalence with the norm $\Psi(w(t))$ introduced in (155):

$$V(t) \leq \frac{ce^c}{2}\Psi(w(t))\,, \tag{177}$$
$$\Psi(w(t)) \leq 8V(t)\,. \tag{178}$$

We have to introduce the factor c on V_1, V_2 to achieve an upper bound of the form $\dot{V} \leq -f(c)V(t)$ where $f(c)>0$. Indeed one can take a look at (121) and see that there is no c factor on the terms in $w^2(0,t)$.

Proof: [Proof of Lemma 8.] We begin with the assumption—to be enforced shortly with a restriction on the initial condition—that $w^2(0,t)\leq B_\nu^2\,,\ t\geq 0$. Since

$$w^2(0,t) \leq 8V(t)\,, \tag{179}$$

if we chose $V(0)\leq \frac{B_\nu^2}{8}$ and prove that V is decreasing, the assumption is validated for all $t\geq 0$. We need to adapt the upper bounds (82)-(92) as well as introducing new ones.

With Lemmas 7 and 1, when $|w(0,t)|\leq B_\nu$ we have

$$\|l(\cdot,w(0,t))\|_\infty \leq \bar{l}\,, \tag{180}$$
$$\|l_\nu(\cdot,w(0,t))\|_\infty \leq \bar{l}_\nu\,, \tag{181}$$
$$\|k_\nu(\cdot,w(0,t))\|_\infty \leq \bar{k}_\nu\,, \tag{182}$$
$$\|k_{x\nu}(\cdot,w(0,t))\|_\infty \leq \bar{k}_{x\nu}\,, \tag{183}$$
$$\|\tilde{k}(\cdot,w(0,t))\|_\infty \leq \epsilon\,, \tag{184}$$
$$\|\tilde{k}_\nu(\cdot,w(0,t))\|_\infty \leq \epsilon\,, \tag{185}$$
$$\|\Omega(\cdot,t)\|_\infty \leq \bar{\Omega}\int_0^1 |w(y,t)|dy\,, \tag{186}$$
$$\begin{aligned} |\Omega_x(x,t)| &\leq \bar{\Omega}_{x1}\int_0^1 |w(y,t)|dy \\ &\quad + \bar{k}_\nu|w(x,t)|\,, \end{aligned} \tag{187}$$
$$|\Gamma(t)| \leq \epsilon\bar{\Gamma}\int_0^1 |w(y,t)|dy\,, \tag{188}$$
$$\begin{aligned} |\Gamma_t(t)| &\leq \epsilon\bar{\Gamma}_1|w_x(0,t)|\int_0^1 |w(y,t)|dy \\ &\quad + \epsilon\bar{\Gamma}\int_0^1 |w_x(y,t)|dy\,, \end{aligned} \tag{189}$$

where $\bar{l}, \bar{l}_\nu$ are defined in Lemma 3 and $\bar{k}_\nu, \bar{k}_{x\nu}$ are defined in Lemma 1 and $\bar{\Omega} := \bar{k}_\nu(1+\bar{l}), \bar{\Omega}_{x1} := \bar{k}_\nu\bar{l} + \bar{k}_{x\nu}(1+\bar{l}), \bar{\Gamma} := 1+\bar{l}, \bar{\Gamma}_1 := 1+\bar{l}+\bar{l}_\nu+\bar{\Omega}(1+\bar{l})$. The proofs of (186) and (187) are almost identical as the ones for (91) and (92) and so we focus on proving (188) and (189).

<u>For Γ:</u> Using (184), (180) and the triangular inequality we have the desired result.

<u>For Γ_t:</u> Taking the derivative of (152) we have that

$$\begin{aligned} -\Gamma_t(t) &= w_t(0,t)\int_0^1 \tilde{k}_\nu(1-y,w(0,t)) \times \\ &\quad \Big[w(y,t) + \int_0^y l(y-s,w(0,t))w(s,t)ds\Big]dy \\ &\quad + \int_0^1 \tilde{k}(1-y,w(0,t)) \\ &\quad \times \Big[w_x(y,t) - w_x(0,t)\Omega(y,t) \\ &\quad + w_t(0,t)\int_0^y l_\nu(y-s,w(0,t))w(s,t)ds\Big]dy \\ &\quad + \int_0^1 \tilde{k}(1-y,w(0,t))\int_0^y l(y-s,w(0,t)) \end{aligned}$$

$$\times \left[w_x(s,t) - w_x(0,t)\Omega(s,t) \right] dsdy\,. \tag{190}$$

Since $w_t(0,t) = w_x(0,t)$, using (187), (184), (185), (180), (181) and (186) we arrive at (189). Also let's notice that Cauchy-Schwarz's inequality provide these very useful upper bounds

$$\int_0^1 |w(x,t)|dx \leq \sqrt{\int_0^1 w^2(x,t)dx}\,, \tag{191}$$

$$\int_0^1 |w_x(x,t)|dx \leq \sqrt{\int_0^1 w_x^2(x,t)dx}\,, \tag{192}$$

$$\forall t \geq 0\,.$$

Let's compute the derivative of V for $t \geq 0$.

Estimate of V_3: Taking the derivative of (174), using (6), $c \geq 1$ and Young's inequality we have that

$$\dot{V}_3(t) \leq \frac{cw^2(0,t)}{4} + \frac{cw_x^2(0,t)}{16}\,. \tag{193}$$

Estimate of V_1: Taking the derivative of (172), we have that

$$\dot{V}_1 = I_{11} + I_{12}\,, \tag{194}$$

$$I_{11}(t) = c\int_0^1 e^{cx}w(x,t)w_x(x,t)dx\,, \tag{195}$$

$$I_{12}(t) = -cw_x(0,t)\int_0^1 e^{cx}w(x,t)\Omega(x,t)dx\,. \tag{196}$$

We first work on I_{11}. Integrating by parts,,

$$\begin{aligned} I_{11}(t) &= \frac{ce^c}{2}\Gamma^2(t) - \frac{c}{2}w^2(0,t) - cV_1(t) \\ &\leq \epsilon^2\bar{\Gamma}^2 e^c V_1(t) - \frac{c}{2}w^2(0,t) - cV_1(t)\,. \end{aligned} \tag{197}$$

The bound (197) follows from (188) and (191). For I_{12} we use (186), (191) as well as Young's inequality. We then have

$$I_{12}(t) \leq \frac{c}{8}w_x^2(0,t) + 4\bar{\Omega}^2 e^c V_1(t)\|w(t)\|^2 \tag{198}$$

Gathering (197) and (198) we have

$$\begin{aligned} \dot{V}_1(t) \leq\ & \frac{c}{8}w_x^2(0,t) - \frac{c}{2}w^2(0,t) - V_1(t) \\ & \times \left[c - \epsilon^2\bar{\Gamma}^2 e^c - 4\bar{\Omega}^2 e^c\|w(t)\|^2 \right]\,. \end{aligned} \tag{199}$$

Estimate of V_2: Taking the derivative with respect to x of (6) and using (150) gives the following system satisfied by w_x

$$w_{xt}(x,t) = w_{xx}(x,t) - w_x(0,t)\Omega_x(x,t)\,, \tag{200}$$

$$\begin{aligned} w_x(1,t) &= \Gamma_t(t) + w_x(0,t)\Omega(1,t)\,, \\ & \forall (x,t) \in [0,1]\times\mathbb{R}^+\,. \end{aligned} \tag{201}$$

We can then take the derivative of V_2 (defined in (173)). We then have

$$\dot{V}_2 = I_{21} + I_{22}\,, \tag{202}$$

$$I_{21}(t) = c\int_0^1 e^{cx}w_x(x,t)w_{xx}(x,t)dx\,, \tag{203}$$

$$I_{22}(t) = -cw_x(0,t)\int_0^1 e^{cx}w_x(x,t)\Omega_x(x,t)dx\,. \tag{204}$$

We first work I_{21}. Integrating by parts,

$$I_{21}(t) = \frac{ce^c}{2}w_x^2(1,t) - \frac{c}{2}w_x^2(0,t) - cV_2(t) \tag{205}$$

Using (201), (189), (187), (192) and (191) we have that

$$\begin{aligned} |w_x(1,t)| \leq\ & \epsilon\bar{\Gamma}\|w_x(t)\| + \epsilon\bar{\Gamma}_1|w_x(0,t)|.\|w(t)\| \\ & + \bar{\Omega}|w_x(0,t)|.\|w(t)\|, \quad t \geq 0\,. \end{aligned} \tag{206}$$

We thus have for $t \geq 0$,

$$\begin{aligned} I_{21}(t) \leq & e^c\Big[4\epsilon^2\bar{\Gamma}^2 V_2(t) + 4\epsilon^2 w_x^2(0,t)V_1(t) \\ & + 2\bar{\Omega}^2 w_x^2(0,t)V_1(t) \Big] - \frac{c}{2}w_x^2(0,t) - cV_2(t)\,. \end{aligned} \tag{207}$$

Moving to I_{22}, using (187), Cauchy-Schwarz and Young's inequalities we have that

$$I_{22}(t) \leq \frac{cw_x^2(0,t)}{4} + 4e^c V_2(t)\|w(t)\|^2(\bar{k}_\nu^2 + \bar{\Omega}_{x1}^2)\,. \tag{208}$$

Gathering (207), (208) we have that

$$\begin{aligned} \dot{V}_2(t) \leq & -w_x^2(0,t)\left[\frac{c}{4} - 2e^c\epsilon^2 V(t)(2+\bar{\Omega}^2) \right] \\ & - \left[c - 4e^c(\epsilon^2\bar{\Gamma}^2 + 2V(t)(\bar{k}_\nu^2 + \bar{\Omega}^2)) \right] V_2(t)\,. \end{aligned} \tag{209}$$

noticing that $V_1 \leq V$, $\|w(t)\|^2 \leq 2V(t)$.

Estimate of V: Gathering (199), (209) and (193) we have

$$\dot{V}(t) \leq -w_x^2(0,t)\left[\frac{c}{16} - 2e^c\epsilon^2 V(t)(2+\bar{\Omega}^2) \right] \tag{210}$$

$$- V_1(t)\left[c - \epsilon^2\bar{\Gamma}^2 e^c - 8\bar{\Omega}^2 e^c V(t) \right] \tag{211}$$

$$- V_2(t)\Big[c - 4e^c(\epsilon^2\bar{\Gamma}^2 + 2V(t)(\bar{k}_\nu^2 + \bar{\Omega}^2)) \Big] \tag{212}$$

$$- 2cV_3(t)\,. \tag{213}$$

Considering (211), we introduce

$$R_1 := \frac{ce^{-c}}{32\bar{\Omega}^2} > 0\,, \tag{214}$$

$$\epsilon_1 := \frac{e^{-\frac{c}{2}}}{\bar{\Gamma}}\sqrt{\frac{c}{2} - 8\bar{\Omega}^2 e^c R_1} > 0\,. \tag{215}$$

That way if we choose $\epsilon \leq \epsilon_1$, $V(0) \leq R_1$, we would have (211) upper bounded by $-\frac{c}{2}V_1(t)$. We then consider (210), and introduce

$$\epsilon_2 := \min\left\{ \epsilon_1, e^{-\frac{c}{2}}\sqrt{\frac{c}{32R_1(2+\bar{\Omega}^2)}} \right\} > 0\,, \tag{216}$$

That way if we choose $\epsilon \leq \epsilon_1$, $V(0) \leq R_1$ we would have (210) negative and (211) bounded by $-\frac{c}{2}V_1(t)$. Finally we consider (212) and introduce

$$R_2 := \min\left\{ R_1, \frac{ce^{-c}}{32(\bar{k}_\nu^2 + \bar{\Omega}_{x1}^2)} \right\} > 0\,, \tag{217}$$

$$\epsilon^* := \min\left\{ \epsilon_2, \frac{e^{\frac{-c}{2}}}{2\bar{\Gamma}} sqrt\frac{c}{2} - 8e^c R_2(\bar{k}_\nu^2 + \bar{\Omega}_{x1}^2) \right\} > 0\,. \tag{218}$$

That way if we choose $\epsilon \leq \epsilon^*$, $V(0) \leq R_2$ we would have (210) ≤ 0, (211) $\leq -\frac{c}{2}V_1(t)$ and (212) $\leq -\frac{c}{2}V_2(t)$. We make the final choice $\epsilon \leq \epsilon^*$ and

$$V(0) \leq \min\left\{ R_2, \frac{B_\nu^2}{8} \right\} =: R_0 > 0\,,\,, \tag{219}$$

where the last condition in (219) is made to ensure that $w^2(0,t) \le B_\nu^2$. With this choice we have that

$$\dot{V}(t) \le -\frac{c}{2}V(t),\ \forall t \ge 0\,. \tag{220}$$

With the comparison principle we complete the proof. ■

Lemma 9: For all $c \ge 1$ there exist functions $\epsilon^*(\breve{B}, c) = \mathcal{O}_{c\to\infty}(e^{-\frac{c}{2}}), \Psi_0(B_\nu, \breve{B}, c) = \mathcal{O}_{c\to\infty}(e^{-2c})$ with a decreasing dependence in the argument $\breve{B}$ such that, for any $\beta \in H$ and any NO approximating k, k_ν with any $\epsilon \in (0, \epsilon^*)$, where ϵ^* is defined in (218), if the perturbed target system (150), (149) is initialized with $w_0 := w(.,0)$ such that $\Psi(w_0) \le \Psi_0$, then

$$\Psi(t) \le 4ce^c\Psi(0)e^{-\frac{c}{2}t}, \quad t \ge 0\,, \tag{221}$$

where

$$\Psi(w(t)) = \|w(t)\|^2 + \|w_x(t)\|^2 + w^2(0,t)\,. \tag{222}$$

Proof: We just have to notice that the Lyapunov function defined in (171) satisfies $V(t) \le \frac{ce^c}{2}\Psi(w(t))$ and $\Psi(w(t)) \le 8V(t)$. We then use Lemma 9, keeping ϵ^*, and introducing $\Psi_0 := \frac{2e^{-c}}{c}R_0$, where R_0 is defined in Lemma. 9 ■

C. Local Stability in Original Plant Variable

We come back to system (1), (2), working with the norm

$$\Omega(u(t)) := \|u(t)\|^2 + \|u_x(t)\|^2 + u^2(0,t) \tag{223}$$

and state the norm equivalence between the perturbed target system (149), (7) and the original system (1), (2).

Lemma 10: [Equiv. norm perturbed-target system] There exist $\rho, \delta > 0$ such that, $\forall t \ge 0$, if $|u(0,t)| \le B_\nu$ then,

$$\Psi(w(t)) \quad \le \quad \delta\,\Omega(u(t))\,, \tag{224}$$

$$\Omega(u(t)) \quad \le \quad \rho\,\Psi(w(t))\,. \tag{225}$$

Proof: Noticing that $w(0,t) = u(0,t)$, we adapt the proof of Lemma 6 and choose $\delta := (6 + 4\bar{k}^2 + 4\bar{k}_x^2)$ and $\rho := (6 + 4\bar{l}^2 + 4\bar{l}^2)$, where $\bar{l}$ is defined in Lemma 7 and $\bar{k} := B_\beta e^{B_\beta}, \bar{k}_x := B_{\beta_x}(1+\bar{k}) + B_\beta \bar{k}, \bar{l}_x := \bar{k}_x + \bar{l}\bar{k} + \bar{k}_x\bar{l}$. ■

With all the lemmas, we now prove Theorem 5.

Proof: [Theorem 5] Adapting the proof of Thm. 3, choose

$$\Omega_0 \quad := \quad \min\left\{B_\nu^2, \frac{\Psi_0}{\delta}\right\}, \tag{226}$$

$$M \quad := \quad \rho\delta c e^c\,, \tag{227}$$

where ρ, δ are defined in Lemma 10 and Ψ_0 defined in Lemma 9. We complete the proof with ϵ^* as defined in Lemma 9. ■

IX. Simulations

We consider the simulation of Eq (1), (2) where the recirculation function, β, is defined as $\beta(x, u(0,t)) = 5\cos((\gamma + u(0,t))\cos^{-1}(x))$ representing a Chebyshev polynomial of the first kind. Note that the $\gamma + u(0,t)$ term controls the resulting shape of the polynomial where $\gamma + u(0,t) = 1$ is a line, $\gamma + u(0,t) = 2$ is a parabola and so on. To simulate the PDE, we use a traditional first-order finite difference upwind scheme with temporal step $dt = 1e{-}4$ and spatial step $dx = 1e{-}2$. In Figure 1, we present the resulting PDE with open-loop control ($U = 0$) and note that the dynamics result in a limit cycle.

Spatial Step Size (dx)	Analytical Kernel Calculation Time(s) ↓	Neural Operator Kernel Calculation Time(s) ↓	Speedup ↑
0.01	0.043	0.028	1.53x
0.001	2.6	0.029	90x
0.0005	10	0.030	333x
0.0001	245	0.057	4298x

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

Then, in Fig. 2, we demonstrate the region of attraction of the proposed gain scheduling controller when implemented with the analytical kernel. We emphasize that, as theoretically shown, gain scheduling is locally stabilizing, but as shown in Fig. 3, it provides better control in contrast to the failure of a linear backstepping controller designed for a linearization at the origin. In particular, Fig. 3 explicitly demonstrates the poor performance of the purely linear controller $U(t) = \int_0^x k(1-y,0)u(y,t)dy$ as it fails to control the system on the right while performing poorly when compared with the gain-scheduled controlled for the system on the left. Thus, Figures 2 and 3 emphasize the known results of [25] and [9]. However, the authors reiterate that these simulations are extremely expensive to simulate due to the recalculation of the analytical kernel at every timestep. Thus, there is significant need for computational speedup to apply the methodology for any real-world application.

We now discuss the training process for the neural operator approximation of the kernel function. To effectively train the kernel operator, one must delicately select a diverse dataset in $\beta(x,\nu), x \in [0,1], \nu \in \mathbb{R}$ in order to have significant coverage for the possible $u(0,t)$ boundary values as the system evolves in time. To effectively build a diverse dataset, we jointly sample 100 γ values from $\gamma \sim$ Uniform$(3,8)$ and for each one of those γ values, we sample 200 ν values from $\nu \sim$ Uniform$(-5,5)$ to create a total dataset size of 20000 points. We emphasize that this large dataset is needed for sufficient coverage across $u(0,t)$ and if one plans to use neural operator approximations with larger initial condition, then one needs to expand their range of ν and most likely needs to increase the number of ν samples per γ value.

Additionally, as theoretically discussed in both Section VIII and IV, the operator mapping requires the approximation of the derivatives with respect to both x, ν and the mixed double derivative with respect to both arguments. In practice, we found that the resulting neural operator kernel is best trained to just approximate the mapping from $\beta(x,\nu) \mapsto k$ without the derivatives. This is for two reasons. First, from a training, implementation, and speedup perspective, the simpler the mapping for computational calculation, the better the performance. For example, the gain-only approach in Section VIII would double the network output size to include calculation of the derivative $\frac{\partial \mathcal{K}}{\partial \nu}$ and the full operator approach in Section IV would quadruple the output layer size to include all the derivatives. Second, from a performance perspective, we

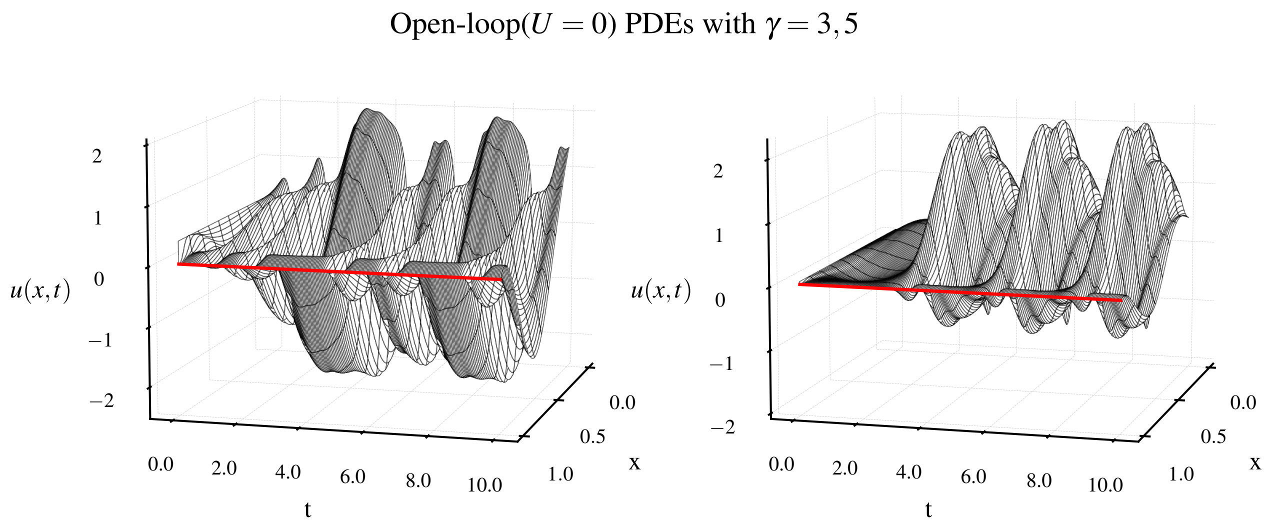

Fig. 1. Open-loop (U=0) simulation of the PDE with the modified Chebyshev polynomial functions $\boldsymbol{\beta(x, u(0,t)) = 5\cos((\gamma + u(0,t))\cos^{-1}(x))}$ with initial conditions $\boldsymbol{u(x,0) = 0.38, 0.04}$ and parameters $\boldsymbol{\gamma = 3, 5}$ for the let and right images respectively.

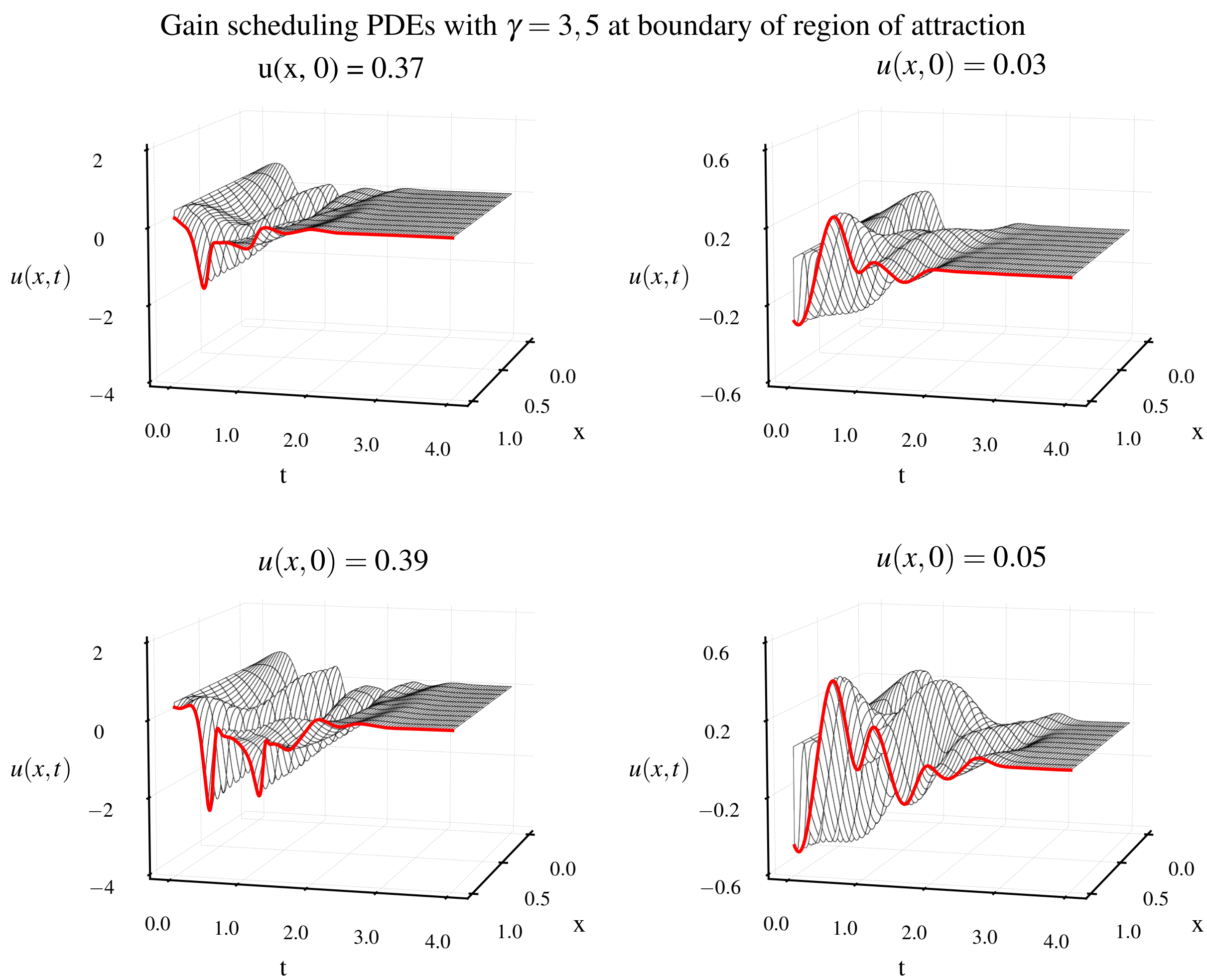

Fig. 2. Analytical solution with gain scheduling for the modified Chebyshev polynomial functions $\boldsymbol{\beta(x, \mu) = 5\cos((\gamma + \mu)\cos^{-1}(x))}$ with parameters $\boldsymbol{\gamma = 3, 5}$ for the left and right images respectively. The top row shows initial conditions 0.37 (left) and 0.03 (right) respectively and the bottom row shows increased initial conditions of 0.39 (left) and 0.05(right). Naturally, the PDE becomes harder to stabilize and for larger initial conditions, gain scheduling fails to control the PDE.

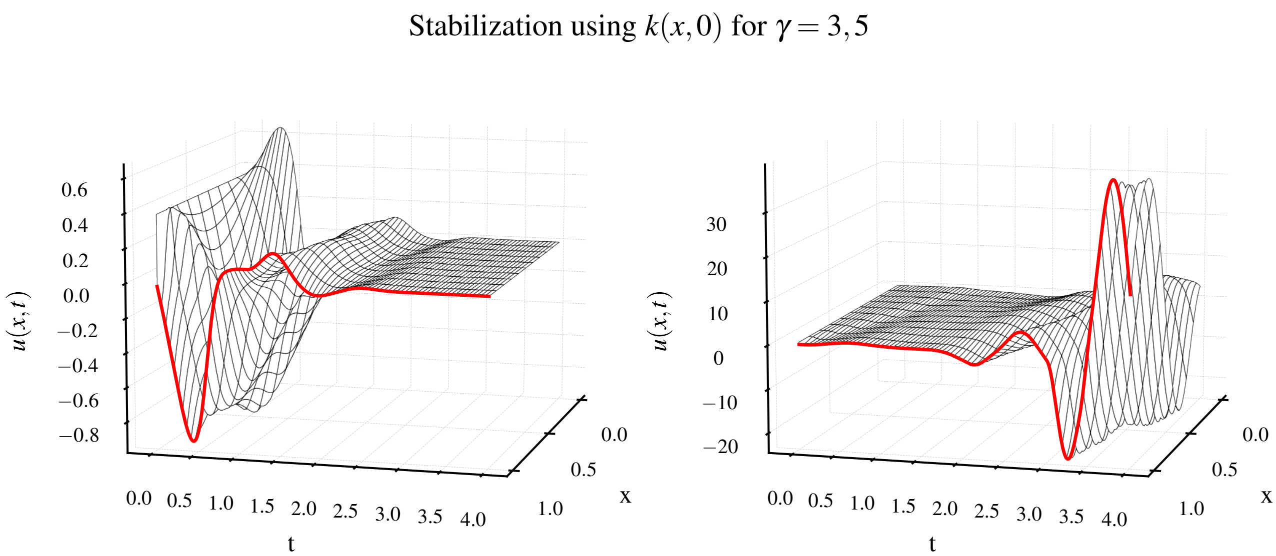

Fig. 3. PDE stabilization using a controller based on a linearization at the origin, with no gain scheduling, $U(t) = \int_0^x k(1-y,0)u(y,t)dy$, for the PDE's corresponding to Figure 1

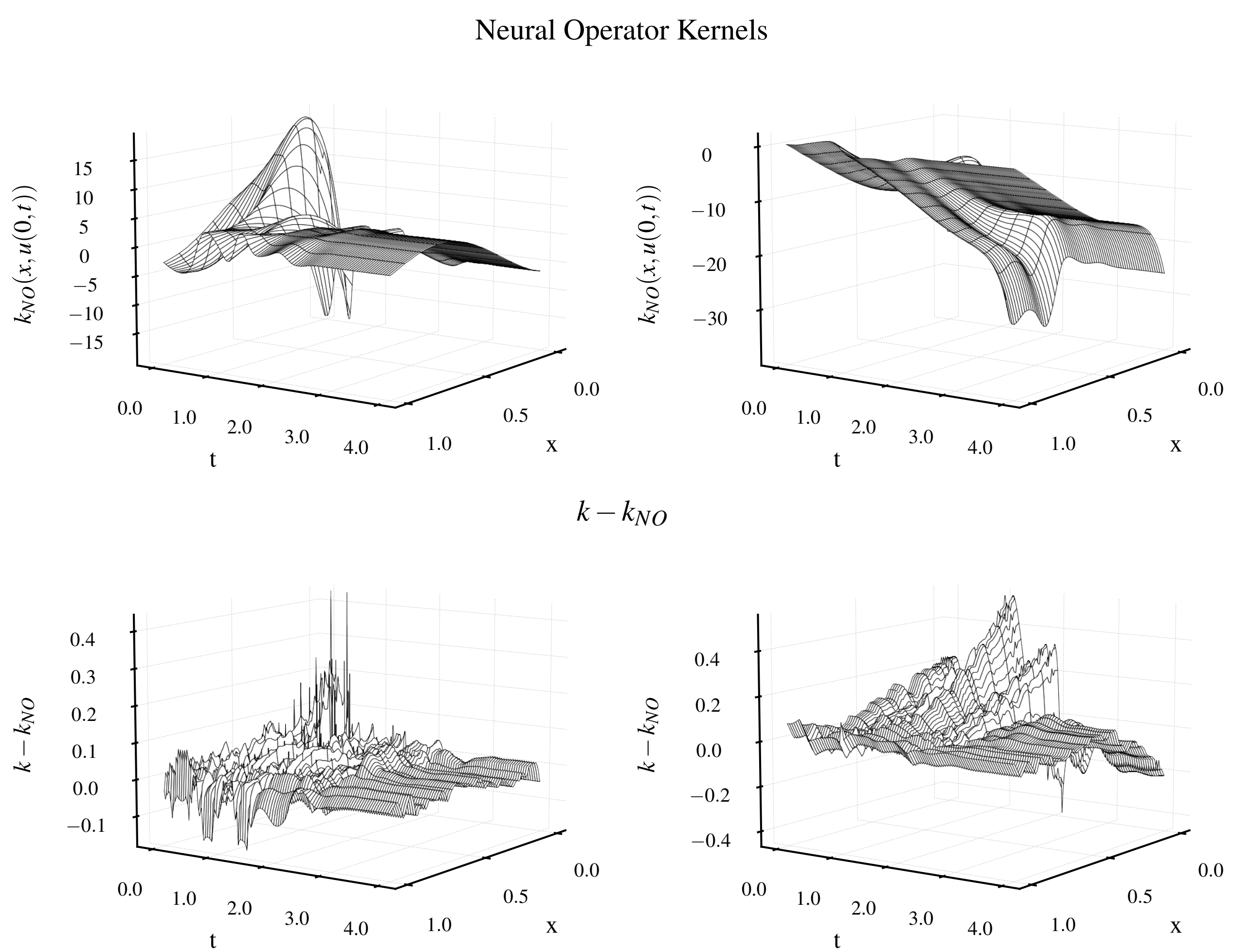

Fig. 4. Neural operator kernels when controlling the PDE in Figure 1(top row), and the difference in kernel error between the resulting gain scheduling kernels (bottom row).

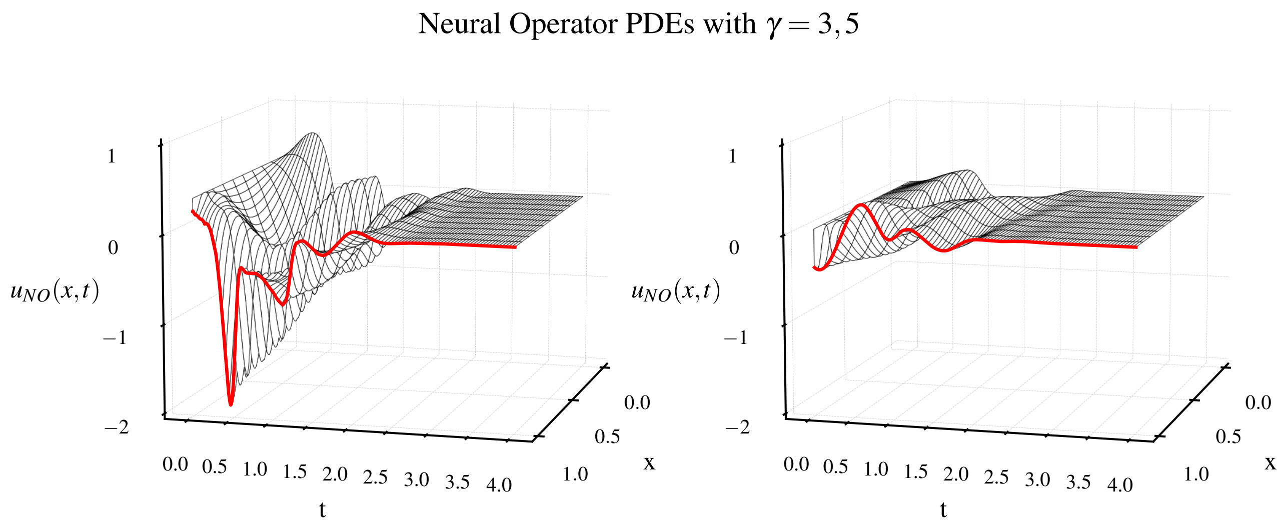

Fig. 5. Stabilization of the plants in Figure 1 with the neural operator approximated kernels in the gain scheduling feedback law.

found that including the derivatives significantly worsened the computational performance of the $\mathcal{K}$ approximation. This is due to the large difference in derivative magnitudes (values can range from 10^3 to 10^{-5} for various combinations in the given dataset) dominating the loss function causing the optimizer to favor accurately approximating the derivative over the operator solution $\mathcal{K}$ even when standard normalization techniques are applied. Furthermore, the authors further considered an alternative physics informed neural network (PINN) approach [29], [21], [17] where the mapping learned is still $\beta(x, \nu) \mapsto k$, but the loss function includes a weighted penalty term that penalizes when the $\mathcal{K}$ derivatives are different from the analytical solution using a 5 point finite difference stenciling [1]. After tuning the penalty weighting, this approach leads to better results than learning the entire operator, but the increase in training time (almost 2x) is expensive and the error on just $\mathcal{K}$ is still larger then learning the mapping without any derivative penalty terms.

The training of the neural operator uses a modification of the DeepXDE package [19] and takes approximately 200 seconds to train on a Nvidia RTX 3090. The neural operator trained has $282,625$ parameters with the branch net as a traditional 4-layer multi-layer perception (MLP) network and the trunk net containing a 2-layer MLP. For ease of future research, all of the hyperparameter details are readily available on Github in an accessible Jupyter notebook. The kernel L_2 training error was 1.96×10^{-3} and the L_2 testing error was 2.04×10^{-3}. Figure 4 shows both the analytical kernel (top row), learned kernel (middle row), and the error (bottom row) with a maximum of approximately 10% between the analytical and learned kernel when applied to the plant in 1. However, this error becomes very small when taken in the broader control feedback loop. For example, in Figure 5, one can see that the closed-loop feedback control with the neural operator approximated kernel effectively stabilizes both systems (top row) with minimal error when compared to stabilization with the analytical kernel (bottom row). Lastly, we conclude the discussion of neural operator approximated simulations by presenting the neural operator approximation performance speedups for various step sizes in Table IX. As expected, the neural operator performance gained scales as the number of discretization points increases (spatial step reduces) and for steps sizes of $dx = 10^{-3}, 10^{-4}$, speedups are on magnitudes of $10^2, 10^3$ resulting in a reduction of *4 mins to 0.5 seconds* for kernel calculation. We emphasize that this is for a single kernel calculation which needs to be completed *at every single timestep* and thus every speedup enhancement gets compounded as the time horizon increases.

X. Conclusion

In this paper, we capitalize on the initial framework in [4] for gain-scheduling of nonlinear hyperbolic PDEs. We introduce the gain-schedulable kernel operator and show that is can be accurately approximated, both theoretically and numerically, by a DeepONet. Then, we present an H_1-norm Lyapunov analysis to prove local stability of the resulting gain-scheduled controller. We also present a similar result for the ”gain-only” approach where only the boundary component of the gain kernel is approximated. We conclude by showcasing simulations of a stabilizing gain-scheduled control law for two challenging hyperbolic PDE problems with nonlinear Chebyshev recirculating coefficients. The resulting simulations demonstrate the operator approximated gain functions achieve numerical speedups of the order of magnitude of 10^3 *at every single timestep in the calculation* paving the way for real-time implementation of gain scheduling feedback laws.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1964.

[2] A. Banach and W. Baumann, “Gain-scheduled control of nonlinear partial differential equations,” in *29th IEEE Conference on Decision and Control*, 1990, pp. 387–392 vol.2.

[3] N. Bekiaris-Liberis and R. Vazquez, “Nonlinear bilateral output-feedback control for a class of viscous Hamilton–Jacobi PDEs,” *Automatica*, vol. 101, pp. 223–231, 2019.

[4] L. Bhan, Y. Shi, and M. Krstic, “Neural operators for bypassing gain and control computations in PDE backstepping,” *IEEE Transactions on Automatic Control*, pp. 1–16, 2023.

[5] T. Chen and H. Chen, “Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems,” *IEEE Transactions on Neural Networks*, vol. 6, no. 4, pp. 911–917, 1995.

[6] J. Coron, R. Vazquez, M. Krstic, and G. Bastin, “Local exponential H^2 stabilization of a 2×2 quasilinear hyperbolic system using backstepping,” *SIAM Journal on Control and Optimization*, vol. 51, no. 3, pp. 2005–2035, 2013.

[7] B. Deng, Y. Shin, L. Lu, Z. Zhang, and G. E. Karniadakis, “Approximation rates of deeponets for learning operators arising from advection–diffusion equations,” *Neural Networks*, vol. 153, pp. 411–426, 2022.

[8] E. Fernández-Cara and E. Zuazua, “Null and approximate controllability for weakly blowing up semilinear heat equations,” *Annales de l’I.H.P. Analyse non linéaire*, vol. 17, no. 5, pp. 583–616, 2000.

[9] I. Karafyllis, N. Espitia, and M. Krstic, “Event-triggered gain scheduling of reaction-diffusion PDEs,” *SIAM Journal on Control and Optimization*, vol. 59, no. 3, pp. 2047–2067, 2021.

[10] I. Karafyllis and M. Krstic, “Global stabilization of a class of nonlinear reaction-diffusion partial differential equations by boundary feedback,” *SIAM Journal on Control and Optimization*, vol. 57, no. 6, pp. 3723–3748, 2019.

[11] ——, “Spill-free transfer and stabilization of viscous liquid,” *IEEE Transactions on Automatic Control*, vol. 67, no. 9, pp. 4585–4597, 2022.

[12] M. Krstic and A. Smyshlyaev, “Backstepping boundary control for first-order hyperbolic PDEs and application to systems with actuator and sensor delays,” *Systems & Control Letters*, vol. 57, no. 9, pp. 750–758, 2008.

[13] ——, *Boundary Control of PDEs: A Course on Backstepping Designs*. SIAM, 2008.

[14] M. Krstic, L. Bhan, and Y. Shi, “Neural operators of backstepping controller and observer gain functions for reaction-diffusion PDEs,” 2023, arXiv:2303.10506.

[15] M. Krstic, L. Magnis, and R. Vazquez, “Nonlinear stabilization of shock-like unstable equilibria in the viscous Burgers PDE,” *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1678–1683, 2008.

[16] S. Lanthaler, Z. Li, and A. M. Stuart, “The nonlocal neural operator: Universal approximation,” 2023, arXiv:2304.13221.

[17] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, K. Azizzadenesheli, and A. Anandkumar, “Physics-informed neural operator for learning partial differential equations,” 2023, arXiv:2111.03794.

[18] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, “Learning nonlinear operators via deeponet based on the universal approximation theorem of operators,” *Nature Machine Intelligence*, vol. 3, no. 3, pp. 218–229, 2021.

[19] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, “DeepXDE: A deep learning library for solving differential equations,” *SIAM Review*, vol. 63, no. 1, pp. 208–228, 2021.

[20] J. Qi, J. Zhang, and M. Krstic, “Neural operators for delay-compensating control of hyperbolic PIDEs,” 2023, arXiv:2307.11436.

[21] M. Raissi, P. Perdikaris, and G. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[22] W. J. Rugh and J. S. Shamma, “Research on gain scheduling,” *Automatica*, vol. 36, no. 10, pp. 1401–1425, 2000.

[23] W. Rugh, “Analytical framework for gain scheduling,” *IEEE Control Systems Magazine*, vol. 11, no. 1, pp. 79–84, 1991.

[24] J. Shamma and M. Athans, “Analysis of gain scheduled control for nonlinear plants,” *IEEE Transactions on Automatic Control*, vol. 35, no. 8, pp. 898–907, 1990.

[25] A. A. Siranosian, M. Krstic, A. Smyshlyaev, and M. Bement, “Gain scheduling-inspired boundary control for nonlinear partial differential equations,” *Journal of dynamic systems, measurement, and control*, vol. 133, no. 5, p. 12, 2011.

[26] R. Vazquez and M. Krstic, “Control of 1-D parabolic PDEs with Volterra nonlinearities, Part I: Design,” *Automatica*, vol. 44, no. 11, pp. 2778–2790, 2008.

[27] ——, “Control of 1-D parabolic PDEs with Volterra nonlinearities, Part II: Analysis,” *Automatica*, vol. 44, no. 11, pp. 2791–2803, 2008.

[28] S. Wang, M. Diagne, and M. Krstić, “Deep learning of delay-compensated backstepping for reaction-diffusion PDEs,” 2023, arXiv:2308.10501.

[29] S. Wang, H. Wang, and P. Perdikaris, “Learning the solution operator of parametric partial differential equations with physics-informed deeponets,” *Science Advances*, vol. 7, no. 40, p. eabi8605, 2021.

Maxence Lamarque is a graduate student at École 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 Vanderbilt 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.

Rafael Vazquez (Senior Member, IEEE) received the electrical engineering and mathematics degrees from the University of Seville, Spain, and the M.Sc. and Ph.D. degrees in aerospace engineering from the University of California, San Diego. He is currently Professor in the Aerospace Engineering Department of the University of Seville, Spain. His research interests include control theory, estimation, distributed parameter systems, and optimization, with applications to spacecraft and aircraft guidance, navigation and control and space surveillance and awareness. He currently serves as Associate Editor for Automatica and IEEE Control Systems Letters (L-CSS).

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.