

Neural Operators for Boundary Stabilization of Stop-and-Go Traffic

Yihuai Zhang

YZHANG169@CONNECT.HKUST-GZ.EDU.CN

Ruiguo Zhong

RZHONG151@CONNECT.HKUST-GZ.EDU.CN

Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou)

Huan Yu*

HUANYU@UST.HK

Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou)

Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology

Abstract

This paper introduces a novel approach to PDE boundary control design using neural operators to alleviate stop-and-go traffic instabilities. Our framework leverages neural operators to design control strategies for traffic flow systems. The traffic dynamics are described by the Aw-Rascle-Zhang (ARZ) model, which consists of second-order coupled hyperbolic partial differential equations (PDEs). The backstepping method which involves constructing and solving a backstepping control kernel is widely used for boundary control of such PDE systems, but it requires intensive depth of expertise and can be time-consuming. To overcome these challenges, we present two distinct neural operator (NO) learning schemes aimed at stabilizing the traffic PDE system. The first scheme embeds NO-approximated gain kernels within a predefined backstepping controller, while the second one directly learns a boundary control law. The Lyapunov analysis is conducted to evaluate the stability of the NO-approximated gain kernels and control law. It is proved that the NO-based closed-loop system is practical stable under certain approximation accuracy conditions. To validate the efficacy of the proposed approach, simulations are conducted to compare the performance of the two neural operator controllers with a PDE backstepping controller and a Proportional Integral (PI) controller. While the NO-approximated methods exhibit larger errors compared to the backstepping controller, they consistently outperform the PI controller, demonstrating faster computation speeds across all scenarios. This result suggests that neural operators can significantly expedite and simplify the process of obtaining boundary controllers for freeway traffic stabilization systems.

Keywords: Intelligent traffic systems(ITS), Partial differential equations(PDEs), Neural operator, Backstepping control

1. Introduction

Stop-and-go traffic oscillations are a common phenomenon on freeways, causing increased waiting time, fuel consumption and traffic accidents [Belletti et al. \(2015\)](#); [De Palma and Lindsey \(2011\)](#); [Schönhof and Helbing \(2007\)](#). Therefore, it is of interest to eliminate the traffic congestion with boundary control, implemented with ramp metering or varying speed limits. To describe the dynamics of the traffic oscillations, [Aw and Rascle \(2000\)](#); [Zhang \(2002\)](#) introduced the Aw-Rascle-Zhang (ARZ) model consisting of a 2×2 hyperbolic partial differential equations(PDEs) to describe the density and velocity of traffic systems. Backstepping method is well-studied for stabilizing PDEs [Krstic and Smyshlyaev \(2008\)](#); [Vazquez et al. \(2011\)](#); [Anfinson and Aamo \(2019\)](#). Consid-

* corresponding author

ering disturbances and delays of hyperbolic PDEs, [Auriol and Di Meglio \(2020\)](#) adopted a robust output feedback controller based on backstepping to guarantee the robustness of hyperbolic PDEs. With applications to traffic system, [Yu and Krstic \(2022\)](#) firstly applied backstepping control method to traffic system based on ARZ PDE model.

The implementation of the aforementioned control designs involves numerical schemes, rendering the process time-consuming, particularly when addressing the solution of PDEs and constructing backstepping control kernels. Moreover, it necessitates substantial expertise, given the intricacy involved in constructing the functional mapping of the backstepping transformation. Recently, machine learning (ML) methods have emerged as powerful tools for solving complex algebraic and PDEs. Among these, neural operator (NO) presents exciting advancements [Kovachki et al. \(2023\)](#), such as DeepONet [Lu et al. \(2021\)](#), and Fourier Neural Operator (FNO) [Li et al. \(2020\)](#). These neural operators offer distinctive advantages compared to other traditional ML methods due to their simple settings in solving complex problems. It is of great values for solving PDEs and backstepping kernel equations. [Bhan et al. \(2023b\)](#) adopted neural operators for nonlinear adaptive control. The operator learning framework for accelerating nonlinear adaptive control was proposed. In addition, they apply the operator learning method for bypassing gain and control computations in PDE control [Bhan et al. \(2023a\)](#); [Qi et al. \(2024\)](#); [Krstic et al. \(2024\)](#). It also can be used to accelerate PDE observer computations in traffic systems [Shi et al. \(2022\)](#).

Contributions. In this paper, we design novel neural operator controllers for boundary stabilization of ARZ traffic PDE system. Two mappings, NO-approximated gain kernels and NO-approximated control law, are developed to improve the computation speed of gain kernels and control law. The Lyapunov analysis is conducted to prove the practical local stability of closed-loop system with the NO-approximated controllers. The simulations are also conducted to evaluate the performance of the two NO-approximated methods with several model-based controllers. It is shown that the NO-approximated methods accelerate computation speeds of gain kernels and control law and successfully stabilize the traffic system.

2. Background for ARZ traffic control

The macroscopic traffic on a given road can be described by a 2×2 nonlinear hyperbolic PDE using ARZ model [Aw and Rascle \(2000\)](#); [Zhang \(2002\)](#). The model is defined by:

$$\partial_t \rho + \partial_x(\rho v) = 0, \quad (1)$$

$$\partial_t(v - V(\rho)) + v \partial_x(v - V(\rho)) = \frac{V(\rho) - v}{\tau}, \quad (2)$$

$$\rho(0, t) = \frac{q^*}{v(0, t)}, \quad (3)$$

$$v(L, t) = v^* + U(t), \quad (4)$$

where $\rho(x, t)$ is traffic density, $v(x, t)$ denotes traffic velocity. $U(t)$ is the control input at the outlet of the road section which can be implemented by varying speed limit. It controls the traffic velocity at the boundary of the road to regulate the traffic system. And the spatial and time domain are defined on $(x, t) \in [0, L] \times \mathbb{R}^+$. τ is the reaction time related to the drivers' behavior. $V(\rho)$ is the fundamental diagram which describes the relation between the traffic density and velocity. The fundamental diagram can be defined using Greenshield's model $V(\rho) = v_f \left(1 - \left(\frac{\rho}{\rho_m}\right)^\gamma\right)$, where the v_f is the maximum speed for the traffic flow, ρ_m denotes the maximum density and (ρ^*, v^*) are

the equilibrium points of the system. Also, using the fundamental diagram, we get $V(\rho^*) = v^*$, $q^* = \rho^*V(\rho^*)$. The linearized system is

$$\partial_t \tilde{w}(x, t) + \lambda_1 \partial_x \tilde{w}(x, t) = 0, \quad (5)$$

$$\partial_t \tilde{v}(x, t) - \lambda_2 \partial_x \tilde{v}(x, t) = c(x) \tilde{w}(x, t), \quad (6)$$

$$\tilde{w}(0, t) = -r \tilde{v}(0, t), \quad (7)$$

$$\tilde{v}(L, t) = U(t), \quad (8)$$

and the backstepping boundary control law is designed as

$$U(t) = \int_0^L K^w(L, \xi) \tilde{w}(\xi, t) d\xi + \int_0^L K^v(L, \xi) \tilde{v}(\xi, t) d\xi. \quad (9)$$

where $\lambda_1 = v^*$, $\lambda_2 = -\rho^*V'(\rho^*) - v^*$, $c(x) = -\frac{1}{\tau} e^{-\frac{x}{\tau v^*}}$, $r = \frac{-\rho^*V'(\rho^*) - v^*}{v^*}$. The corresponding backstepping transformation is

$$\alpha(x, t) = \tilde{w}(x, t), \quad (10)$$

$$\beta(x, t) = \tilde{v}(x, t) - \int_0^x K^w(x, \xi) \tilde{w}(\xi, t) d\xi - \int_0^x K^v(x, \xi) \tilde{v}(\xi, t) d\xi, \quad (11)$$

where the control gain kernels in (9) defined in the triangular domain $\mathcal{T} = \{(x, \xi) : 0 \leq \xi \leq x \leq L\}$, are obtained using

$$\lambda_2 K_x^w(x, \xi) - \lambda_1 K_\xi^w(x, \xi) = c(x) K^v(x, \xi), \quad (12)$$

$$\lambda_2 K_x^v(x, \xi) + \lambda_2 K_\xi^v(x, \xi) = 0, \quad (13)$$

$$K^w(x, x) = -\frac{c(x)}{\lambda_1 + \lambda_2}, \quad (14)$$

$$K^v(x, 0) = -K^w(x, 0). \quad (15)$$

We have the following theorem for the closed-loop system,

Theorem 1 (Theorem 2, Yu and Krstic (2019)) *The system (1) - (4) with initial conditions $\rho(x, 0)$, $v(x, 0) \in L^2[0, L]$ and boundary controller (9) is locally exponentially stable in L_2 -sense at finite time $t_f = \frac{L}{\lambda_1} + \frac{L}{\lambda_2}$.*

3. Neural operator for approximating backstepping kernels

The neural operator is employed for approximating the operator mapping of functionals. In the section, we introduce the neural operator using DeepONet for the mapping from the characteristic speed λ_2 to kernels $K^w(x, \xi)$, $K^v(x, \xi)$. A neural operator for approximating a nonlinear mapping $\mathcal{G} : \mathcal{U} \mapsto \mathcal{V}$

$$\mathcal{G}_{\mathbb{N}}(\mathbf{u}_m)(y) = \sum_{k=1}^p g^{\mathcal{N}}(\mathbf{u}_m; \vartheta^{(k)}) f^{\mathcal{N}}(y; \theta^{(k)}), \quad (16)$$

where \mathcal{U}, \mathcal{V} are function spaces of continuous functions $u \in \mathcal{U}$, $v \in \mathcal{V}$. u_m is the evaluation of function u at points $x_i = x_1, \dots, x_m$. p is the number of basis components in the target space, $y \in Y$ is the location of the output function $v(y)$ evaluations, and $g^{\mathcal{N}}$, $f^{\mathcal{N}}$ are NNs termed branch and trunk networks.

Theorem 2 (DeepONet universal approximation theorem) *Let $X \subset \mathbb{R}^{d_x}$, $Y \subset \mathbb{R}^{d_y}$ be compact sets of vectors $x \in X$ and $y \in Y$. Let $\mathcal{U} : X \rightarrow \mathcal{U} \subset \mathbb{R}^{d_u}$ and $\mathcal{V} : Y \rightarrow \mathcal{V} \subset \mathbb{R}^{d_v}$ be sets of continuous functions $u(x)$ and $v(y)$, respectively. And let \mathcal{U} be also compact. Assume the operator*

$\mathcal{G}: \mathcal{U} \rightarrow \mathcal{V}$ is continuous. Then, for all $\epsilon > 0$, there exists a $m^*, p^* \in \mathbb{N}$ such that for each $m \geq m^*, p \geq p^*$, there exist $\theta^{(k)}, \vartheta^{(k)}$, neural networks $f^{\mathcal{N}}(\cdot; \theta^{(k)}), g^{\mathcal{N}}(\cdot; \vartheta^{(k)})$, $k = 1, \dots, p$ and $x_j \in X, j = 1, \dots, m$, with corresponding $\mathbf{u}_m = (u(x_1), u(x_2), \dots, u(x_m))^{\top}$, such that

$$|\mathcal{G}(u)(y) - \mathcal{G}_{\mathbb{N}}(\mathbf{u}_m)(y)| < \epsilon, \quad (17)$$

for all functions $u \in \mathcal{U}$ and all values $y \in Y$ of $\mathcal{G}(u) \in \mathcal{V}$. [Bhan et al. \(2023a\)](#); [Chen and Chen \(1995\)](#)

Definition 3 The kernel operator $\mathcal{K}: \mathbb{R}^+ \rightarrow C^1(\mathcal{T}) \times C^1(\mathcal{T})$ is defined by:

$$K^w(x, \xi) := \mathcal{K}^w(\lambda_2)(x, \xi), K^v(x, \xi) := \mathcal{K}^v(\lambda_2)(x, \xi). \quad (18)$$

The kernel operator \mathcal{K} denotes the mapping from the characteristic speed to backstepping transformation kernels. Based on Theorem 2, we have the following lemma of the approximation of neural operator for the kernel equations:

Lemma 4 For all $\epsilon > 0$, there exists a neural operator $\hat{\mathcal{K}}$ that for all $(x, \xi) \in \mathcal{T}$,

$$\sup_{\lambda_2 \in \mathcal{U}} \left\| \hat{\mathcal{K}}(\lambda_2) - \mathcal{K}(\lambda_2) \right\| < \epsilon. \quad (19)$$

Proof The existence, uniqueness of the kernel equations have been proved in [Vazquez et al. \(2011\)](#). So the mapping $\mathcal{K}: \mathbb{R}^+ \rightarrow C^1(\mathcal{T}) \times C^1(\mathcal{T})$ from λ_2 to $K^w(x, \xi), K^v(x, \xi)$ indicated by (12)-(15) and the solution of the kernel equations exists. The neural operator $\hat{\mathcal{K}}$ approximates the backstepping kernels for a given λ_2 and their derivatives in the whole spatial-temporal domain. Using Theorem 2, the maximum value of approximation errors is less than ϵ . \blacksquare

Remark 5 The error between the neural operator and the kernel operator is less than a given constant ϵ . For the partial derivative of the kernels, we also have the continuous operator $\mathcal{M}: \mathbb{R}^+ \rightarrow C^1(\mathcal{T}) \times C^0(\mathcal{T}) \times C^0(\mathcal{T}) \times C^1(\mathcal{T}) \times C^1(\mathcal{T})$

$$\mathcal{M}(\lambda_2)(x, \xi) := (K(x, \xi), \kappa_1(x, \xi), \kappa_2(x, \xi), \kappa_3(x, x), \kappa_4(x)) \quad (20)$$

where $\kappa_1(x, \xi) = \lambda_2 K_x^w(x, \xi) - \lambda_1 K_\xi^w(x, \xi) - c(x)K^v(x, \xi)$, $\kappa_2(x, \xi) = \lambda_2 K_x^v(x, \xi) + \lambda_2 K_\xi^v(x, \xi)$, $\kappa_3(x, x) = K^w(x, x) + \frac{c(x)}{\lambda_1 + \lambda_2}$, and $\kappa_4(x) = K^v(x, 0) + K^w(x, 0)$. And there exists a neural operator $\hat{\mathcal{M}}$ such that for all $(x, \xi) \in \mathcal{T}$

$$\sup_{\lambda_2 \in \mathcal{U}} \left\| \mathcal{M}(\lambda_2)(x, \xi) - \hat{\mathcal{M}}(\lambda_2)(x, \xi) \right\| < \epsilon. \quad (21)$$

So there exists the neural operator $\mathcal{K}(\lambda_2)(x, \xi)$, such that

$$\begin{aligned} \sup_{\lambda_2 \in \mathcal{U}} \left\| \mathcal{K}(\lambda_2)(x, \xi) - \hat{\mathcal{K}}(\lambda_2)(x, \xi) \right\| &+ \sup_{\lambda_2 \in \mathcal{U}} \left\| \partial_x(\mathcal{K}(\lambda_2)(x, \xi) - \hat{\mathcal{K}}(\lambda_2)(x, \xi)) \right\| \\ &+ \sup_{\lambda_2 \in \mathcal{U}} \left\| \partial_\xi(\mathcal{K}(\lambda_2)(x, \xi) - \hat{\mathcal{K}}(\lambda_2)(x, \xi)) \right\| < \epsilon \end{aligned} \quad (22)$$

We then provide the stability analysis of the ARZ traffic system with the NO-approximated kernels. We first start with the approximated kernels and put them into the ARZ system to get the NO-approximated target system. For a given λ_2 , defining the output of the neural operator $\hat{\mathcal{K}}(\lambda_2)(x, \xi)$:

$$\hat{K}^w = \hat{\mathcal{K}}^w(\lambda_2)(x, \xi), \hat{K}^v = \hat{\mathcal{K}}^v(\lambda_2)(x, \xi). \quad (23)$$

For the NO-approximated kernels $\hat{\mathcal{K}}^w, \hat{\mathcal{K}}^v$, we have the following result,

Theorem 6 *The PDE system (5)-(8) is exponential stable under the control law (9) with approximated kernels \hat{K} at initial conditions $\tilde{w}(x, 0), \tilde{v}(x, 0)$, satisfying*

$$\|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq e^{-\eta t} \frac{m_2}{m_1} \|(\tilde{w}(x, 0), \tilde{v}(x, 0))\|_{L^2}^2, \quad (24)$$

where $m_1 > 0, m_2 > 0, a > 0, \eta = \nu - \frac{2a\epsilon(1+L)(2\lambda_2+\lambda_1)}{m_1\lambda_2}$. The kernels are approximated by the neural operator (19) with accuracy ϵ . It also means that the system (1)-(4) is locally exponential stable under the NO-approximated kernels and the system can eventually achieve to its equilibrium.

Proof We define the error for the NO-approximated kernels and backstepping kernels: $\tilde{K}^w(x, \xi) = K^w(x, \xi) - \hat{K}^w(x, \xi)$, $\tilde{K}^v(x, \xi) = K^v(x, \xi) - \hat{K}^v(x, \xi)$, and the backstepping transformation then is turned into:

$$\hat{\alpha}(x, t) = \tilde{w}(x, t), \quad (25)$$

$$\hat{\beta}(x, t) = \tilde{v}(x, t) - \int_0^x \hat{K}^w(x, \xi) \tilde{w}(\xi, t) d\xi - \int_0^x \hat{K}^v(x, \xi) \tilde{v}(\xi, t) d\xi, \quad (26)$$

the corresponding backstepping control law is

$$U(t) = \int_0^L \hat{K}^w(L, \xi) \tilde{w}(\xi, t) d\xi + \int_0^L \hat{K}^v(L, \xi) \tilde{v}(\xi, t) d\xi. \quad (27)$$

Thus we get the target system with the NO-approximated kernels as

$$\partial_t \hat{\alpha}(x, t) + \lambda_1 \partial_x \hat{\alpha}(x, t) = 0, \quad (28)$$

$$\begin{aligned} \partial_t \hat{\beta}(x, t) - \lambda_2 \partial_x \hat{\beta}(x, t) &= \lambda_2 (\tilde{K}^w(x, 0) + \tilde{K}^v(x, 0)) \tilde{v}(0, t) + (\lambda_1 + \lambda_2) \tilde{K}^w(x, x) \tilde{w}(x, t) \\ &\quad + \int_0^x (\lambda_2 \tilde{K}_x^w(x, \xi) + \lambda_1 \tilde{K}_\xi^w(x, \xi)) \tilde{w}(\xi, t) d\xi \\ &\quad + \int_0^x (\lambda_2 \tilde{K}_x^v(x, \xi) + \lambda_2 \tilde{K}_\xi^v(x, \xi)) \tilde{v}(\xi, t) d\xi, \end{aligned} \quad (29)$$

$$\hat{\alpha}(0, t) = -r \hat{\beta}(0, t), \quad (30)$$

$$\hat{\beta}(L, t) = 0. \quad (31)$$

For the target system (28) - (31) with the NO-approximated kernels, we define the Lyapunov candidate as

$$V_k(t) = \int_0^L \frac{e^{-\frac{\nu}{\lambda_1} x}}{\lambda_1} \hat{\alpha}^2(x, t) + a \frac{e^{-\frac{\nu}{\lambda_2} x}}{\lambda_2} \hat{\beta}^2(x, t) dx, \quad (32)$$

where the coefficients ν and a are constants and $\nu > 0, a > 0$. The states of the NO-approximated backstepping target system $(\hat{\alpha}, \hat{\beta})$ and the original states (\tilde{w}, \tilde{v}) have equivalent L^2 norms, the Lyapunov functional $V_k(t)$ is equivalent to the L^2 norm of the target system, so that there exist two constants $m_1 > 0$ and $m_2 > 0$,

$$m_1 \|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq V_k(t) \leq m_2 \|(\tilde{w}, \tilde{v})\|_{L^2}^2. \quad (33)$$

Taking time derivative along the trajectories of the system, and then we put into the system dynamics, integrating by parts, thus we get

$$\begin{aligned}
 \dot{V}_k(t) &= -\nu V_k(t) + (r^2 - a)\hat{\beta}^2(0, t) - e^{-\frac{\nu}{\lambda_1}L}\hat{\alpha}^2(L, t) \\
 &+ \int_0^L 2a \frac{e^{-\frac{\nu}{\lambda_2}x}}{\lambda_2} \hat{\beta}(x, t) \left(\lambda_2(\tilde{K}^w(x, 0) + \tilde{K}^v(x, 0))\tilde{v}(0, t) + (\lambda_1 + \lambda_2)\tilde{K}^w(x, x)\tilde{w}(x, t) \right. \\
 &+ \int_0^x (\lambda_2\tilde{K}_x^w(x, \xi) + \lambda_1\tilde{K}_\xi^w(x, \xi))\tilde{w}(\xi, t)d\xi \\
 &+ \left. \int_0^x (\lambda_2\tilde{K}_x^v(x, \xi) + \lambda_2\tilde{K}_\xi^v(x, \xi))\tilde{v}(\xi, t)d\xi \right) dx, \tag{34}
 \end{aligned}$$

For the integral term, we take the norm and using the Young inequality and Cauchy-Schwarz inequality, and then combining the equivalent norm of the Lyapunov candidate, we have:

$$\int_0^L \left\| 2a \frac{e^{-\frac{\nu}{\lambda_2}x}}{\lambda_2} \hat{\beta}(x, t) \lambda_2(\tilde{K}^w(x, 0) + \tilde{K}^v(x, 0))\tilde{v}(0, t) \right\| dx \leq \frac{2a\epsilon}{m_1} V_k(t) + 2aL\epsilon\hat{\beta}^2(0, t). \tag{35}$$

Using the same method, we can easily get the results for the other terms of the Lyapunov candidate,

$$\int_0^L \left\| 2a \frac{e^{-\frac{\nu}{\lambda_2}x}}{\lambda_2} \hat{\beta}(x, t) (\lambda_1 + \lambda_2)\tilde{K}^w(x, x)\tilde{w}(x, t) \right\| dx \leq \frac{2a\epsilon(\lambda_1 + \lambda_2)}{m_1\lambda_2} V_k(t), \tag{36}$$

$$\int_0^L \left\| 2a \frac{e^{-\frac{\nu}{\lambda_2}x}}{\lambda_2} \hat{\beta}(x, t) + \int_0^x (\lambda_2\tilde{K}_x^w(x, \xi) + \lambda_1\tilde{K}_\xi^w(x, \xi))\tilde{w}(\xi, t)d\xi \right\| dx \leq \frac{2a\epsilon(\lambda_1 + \lambda_2)L}{m_1\lambda_2} V_k(t), \tag{37}$$

$$\int_0^L \left\| 2a \frac{e^{-\frac{\nu}{\lambda_2}x}}{\lambda_2} \hat{\beta}(x, t) + \int_0^x (\lambda_2\tilde{K}_x^v(x, \xi) + \lambda_2\tilde{K}_\xi^v(x, \xi))\tilde{v}(\xi, t)d\xi \right\| dx \leq \frac{2a\epsilon L}{m_1} V_k(t), \tag{38}$$

thus we get the following Lyapunov candidate,

$$\dot{V}_k(t) \leq -\eta V_k(t) + (r^2 - a + 2aL\epsilon)\hat{\beta}^2(0, t) - e^{-\frac{\nu}{\lambda_1}L}\hat{\alpha}^2(L, t), \tag{39}$$

where $\eta = \nu - \frac{2a\epsilon(1+L)(2\lambda_2+\lambda_1)}{m_1\lambda_2}$. The coefficients ν, ϵ, a are chosen such that

$$\eta > 0, r^2 - a + 2aL\epsilon < 0. \tag{40}$$

So we get the following result:

$$\dot{V}_k(t) \leq -\eta V_k(t) \rightarrow V_k(t) \leq V(0)e^{-\eta t} \tag{41}$$

Using the equivalent norm of the Lyapunov functional, we have:

$$\|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq e^{-\eta t} \frac{m_2}{m_1} \|(\tilde{w}(x, 0), \tilde{v}(x, 0))\|_{L^2}^2. \tag{42}$$

The exponential stability thus is proved. ■

4. NO-approximated backstepping control law

Based on the previous section, we have proved that the PDE system is exponentially stable with the NO-approximated kernels. In this section, we consider whether the neural operator can directly approximate the mapping from the characteristic speed to control law (9) rather than approximating the backstepping kernels. The form of stability we get in this section is also only achieved in the practical sense using the trained neural operator. Different with the results in [Bhan et al. \(2023a\)](#),

we train the mapping only from the characteristic speed λ_2 , while in [Bhan et al. \(2023a\)](#) they trained the mapping from the parameter $\beta(x)$ and the system state $u(x, t)$ at its corresponding time step to the backstepping control law. It costs more training time due to the number of model parameters.

Recalling the control law (9), we define the operator mapping $\mathcal{H}(\lambda_2) : \mathbb{R}^+ \rightarrow \mathbb{R}$ from λ_2 to $U(t)$. From the expression of backstepping control law (9), it is shown that there is no explicit form for the mapping from λ_2 to $U(t)$. The relation between the λ_2 and $U(t)$ is characterized with the kernel equations (12)-(15). The control law for mapping is

$$U(t) = \mathcal{H}(\lambda_2)(L, t), \quad (43)$$

and the NO-approximated mapping for $\mathcal{H}(\lambda_2) : \mathbb{R}^+ \rightarrow \mathbb{R}$ is defined as $\hat{\mathcal{H}}(\lambda_2) : \mathbb{R}^+ \rightarrow \mathbb{R}$. It can be found that the NO-approximated mapping $\hat{\mathcal{H}}(\lambda_2)$ has the following lemma based on Theorem 2,

Lemma 7 *For $\epsilon > 0$, there exists a neural operator $\hat{\mathcal{H}}(\lambda_2)$ that can approximate the control law mapping in the spatial-temporal domain $(x, t) \in [0, L] \times \mathbb{R}^+$:*

$$\left| \mathcal{H}(\lambda_2)(L) - \hat{\mathcal{H}}(\lambda_2)(L) \right| < \epsilon. \quad (44)$$

Applying the NO-approximated control law to the system, we get the target system as:

$$\partial_t \check{\alpha}(x, t) + \lambda_1 \partial_x \check{\alpha}(x, t) = 0, \quad (45)$$

$$\partial_t \check{\beta}(x, t) - \lambda_2 \partial_x \check{\beta}(x, t) = 0, \quad (46)$$

$$\check{\alpha}(0, t) = -r \check{\beta}(0, t), \quad (47)$$

$$\check{\beta}(L, t) = \mathcal{H}(\lambda_2)(L, t) - \hat{\mathcal{H}}(\lambda_2)(L, t). \quad (48)$$

Compared with the target system in [Yu and Krstic \(2019\)](#), the system (45)-(48) is not strictly exponential stable due to the approximated error of NO-control law. We have the following theorem for the NO-approximated control law,

Theorem 8 *The system (5)-(8) is locally practically exponentially stable under the control law (43) with initial conditions $\tilde{w}(x, 0), \tilde{v}(x, 0)$, the kernels are the same as (12)-(15), such that*

$$\|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq \frac{k_2}{k_1} e^{-\nu t} \|(\tilde{w}(x, 0), \tilde{v}(x, 0))\|_{L^2}^2 + \frac{a}{k_1} e^{-\frac{\nu}{\lambda_2} L} \epsilon^2. \quad (49)$$

Proof For the NO-approximated control law, using the Lyapunov candidate again to analyze the stability of the target system (45)-(48).

$$V_U(t) = \int_0^L \frac{e^{-\frac{\nu}{\lambda_1} x}}{\lambda_1} \check{\alpha}^2(x, t) + a \frac{e^{-\frac{\nu}{\lambda_2} x}}{\lambda_2} \check{\beta}^2(x, t) dx, \quad (50)$$

where the coefficients ν and a are the same with before. The Lyapunov functional also has the following equivalent norm with $k_1 > 0, k_2 > 0$

$$k_1 \|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq V_U(t) \leq k_2 \|(\tilde{w}, \tilde{v})\|_{L^2}^2. \quad (51)$$

Taking time derivative along the trajectories, putting into the system dynamics, integrating by parts, we have:

$$\dot{V}_U(t) = -\nu V_U(t) + (r^2 - a) \check{\beta}^2(0, t) + a e^{-\frac{\nu}{\lambda_2} L} \check{\beta}^2(L, t) - e^{-\frac{\nu}{\lambda_1} L} \check{\alpha}^2(L, t). \quad (52)$$

The term $a e^{-\frac{\nu}{\lambda_2} L} \check{\beta}^2(L, t)$ is equal to 0 in the ideal situation when the NO-approximated mapping achieves 100% accuracy of approximation which means that $\mathcal{H}(\lambda_2)(L) - \hat{\mathcal{H}}(\lambda_2)(L) = 0$ and we can easily get the exponential stability for the system. Here the mapping has the error ϵ . We take

the $r^2 - a \leq 0$, and we get

$$V_U(t) \leq V_U(0)e^{-\nu t} + ae^{-\frac{\nu}{\lambda_2}L} \sup_{0 \leq \varsigma \leq t} (\mathcal{H}(\lambda_2)(L) - \hat{\mathcal{H}}(\lambda_2)(L))^2(L, \varsigma). \quad (53)$$

Using (51), we have

$$\|(\tilde{w}, \tilde{v})\|_{L^2}^2 \leq \frac{k_2}{k_1} e^{-\nu t} \|(\tilde{w}(x, 0), \tilde{v}(x, 0))\|_{L^2}^2 + \frac{a}{k_1} e^{-\frac{\nu}{\lambda_2}L} \epsilon^2. \quad (54)$$

Thus we have proved that the system (5)-(8) is locally practically exponentially stable. \blacksquare

5. Experiments

In this section, we present and analyze the performance of the proposed neural operator controllers for the ARZ traffic PDE system, and also provide two comparisons with model-based controllers: (i) Backstepping controller (ii) PI controller.

We first test the efficacy of the controller with *NO-approximated backstepping kernels*. To train the neural operator, we use numerical simulations to get data for training. We run the simulation on a $L = 500\text{m}$ long road and the simulation time is $T = 300\text{s}$. The free-flow velocity is $v_m = 40\text{m/s}$, the maximum density is $\rho_m = 160\text{veh/km}$, the equilibrium density is selected as $\rho^* = 120\text{veh/km}$, the reaction time for the drives adapting to speed is $\tau = 60\text{s}$, and $\gamma = 1$. For the initial conditions, we take the sinusoidal inputs as $\rho(x, 0) = \rho^* + 0.1 \sin\left(\frac{3\pi x}{L}\right) \rho^*$, $v(x, 0) = v^* - 0.1 \sin\left(\frac{3\pi x}{L}\right) v^*$ to mimic the stop-and-go traffic wave. To get enough data for training, we use 900 different values randomly sampled from the $\rho^* \in [90\text{veh/km}, 130\text{veh/km}]$ to get the different value for $\lambda_2 \in [5\text{m/s}, 25\text{m/s}]$ and different kernels $K^w(x, \xi)$, $K^v(x, \xi)$. The corresponding parameters are also calculated using the relations between different parameters. The network structure is the same with the Lu et al. (2021). We adopted the DeepONet framework for the implementation of operator learning. And we train the model on an Nvidia RTX 4090Ti GPU. Using the trained neural operator model, we run the simulation with the same parameters. The results for backstepping controller and NO-approximated controller is shown in Fig. 1 and 2 (a) and (b). The blue line denotes the initial condition while the red line represents the boundary condition of the system. It can be found that the neural operator can still stabilize the traffic system compared with the backstepping control method. The traffic density and velocity all converge to their equilibrium point $\rho^* = 120\text{veh/km}$, $v^* = 36\text{km/h}$. The error between the closed-loop result of the backstepping controller and the approximated one is shown in Fig. 2 (c) and (d).

We then test the *direct NO-approximated controller*. Using the same parameter settings in previous section to get the training data including 900 instance. And we get the result for approximated control law mapping is shown in Fig. 3 (a), (b). We also set the backstepping controller as the baseline to evaluate the performance of the neural operator. The error between the approximated control law and backstepping controller is shown in Fig. 3 (c), (d). It is shown that the density and velocity do not converge to their equilibrium point. The density and velocity error exist in the whole simulation time because we only achieve the practically exponentially stable for the NO-approximated control law in Theorem 8. We also provide a comparison with the PI controller. It was shown that the PI boundary control can also stabilize the traffic system Zhang et al. (2019). We also apply the PI controller at the outlet of the road section as $\tilde{v}(L, t) = U_{PI}(t)$. The control law is given as:

$$U_{PI}(t) = v^* + k_p^v(v(0, t) - v^*) + k_i^v \int_0^t (v(0, t) - v^*) ds, \quad (55)$$

where k_p^v , k_i^v are tuning gains. The results for PI control are shown in Fig. 4. For the control law mapping, we give the comparisons of control law and norm of states, which are shown in Fig. 5.

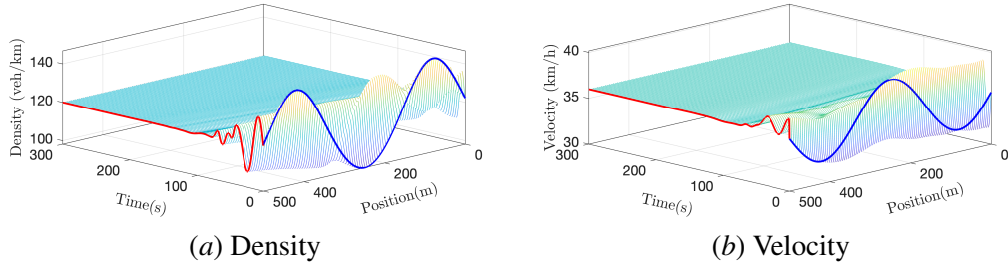
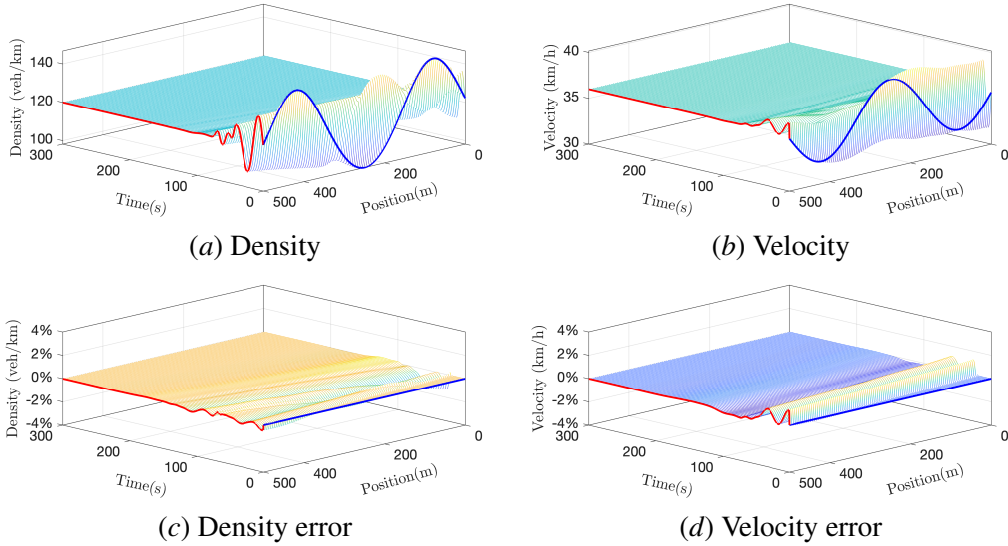


Figure 1: Backstepping controller for closed-loop system


 Figure 2: Neural operator for $\lambda_2 \rightarrow K^w(x, \xi), K^v(x, \xi)$

Method	Average Computation Time(s)	Average L_2 Error
Backstepping controller	5.9×10^{-2}	None
NO-approximated kernels	1.9×10^{-4}	1.23×10^{-3}
NO-approximated control law	1.7×10^{-3}	1.65×10^{-2}
PI controller	1.9×10^{-3}	None

 Table 1: The computation time and average L_2 error of different controllers and neural operators

From the results of $U(t)$, the neural operator can approximate the backstepping kernels and control law which can stabilize the traffic system. All the four controllers are eventually stabilize the system. However, the norm of the states of NO-approximated control law do not converge to zero due to the small oscillations in the simulation period. Overall, it is shown that the NO-approximated kernels and directly approximated control law achieves satisfying closed-loop results. The computation time of neural operators, backstepping controller, PI controller are shown in Tab. 1. We set the backstepping control method as the baseline method of the system. It is can be found that the average computation time of NO-approximated methods are faster than backstepping controller and PI controller which gives the possibility of accelerating the online application in real traffic system with lower relative errors.

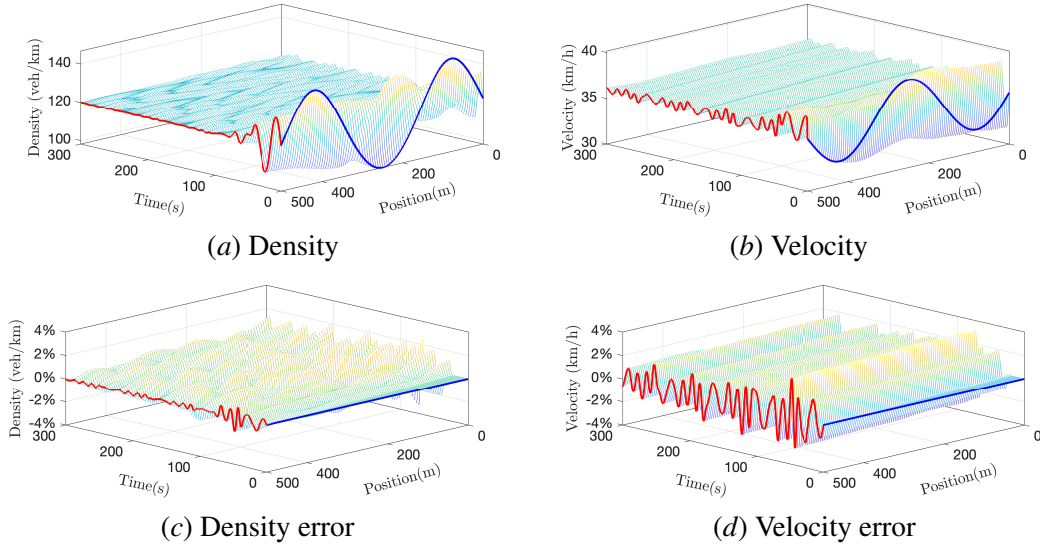
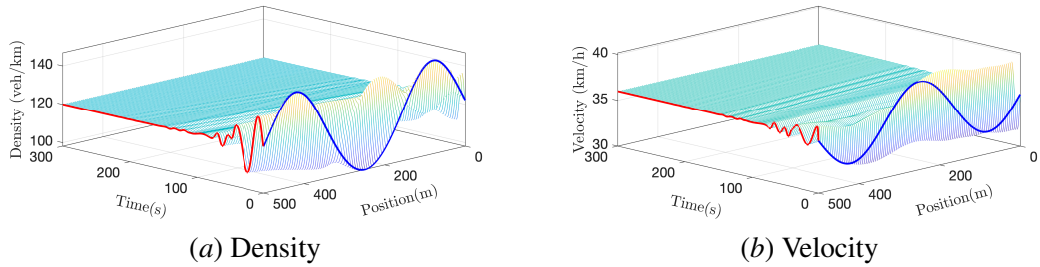
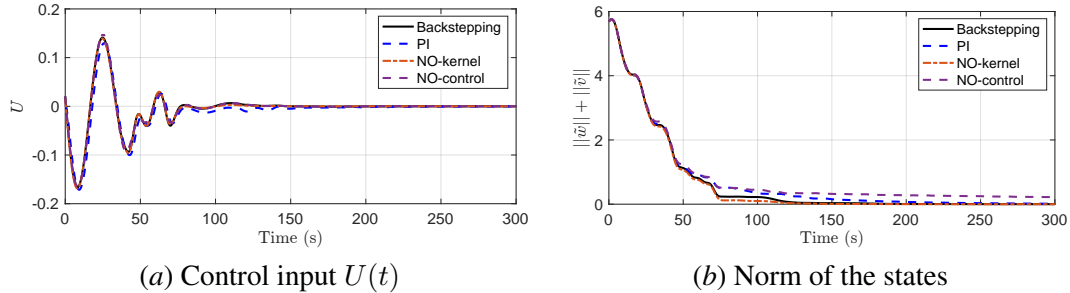

 Figure 3: Neural operator for $\lambda_2 \rightarrow U(t)$


Figure 4: PI controller


 Figure 5: Comparison of $U(t)$ and states norm

6. Conclusion

In this paper, we proposed an operator learning framework for boundary control of traffic systems. The ARZ PDE model is adopted to describe the spatial-temporal evolution of traffic density and velocity. The neural operators using DeepONet are trained to approximate the two mappings. We also derived the theoretical stability for the NO-approximated closed-loop system using Lyapunov analysis. The simulation results showed that the NO-approximated mappings have satisfying accuracy and significantly accelerate the computation process. One of the future work is to incorporate real traffic data into training of neural operator.

References

- Henrik Anfinssen and Ole Morten Aamo. *Adaptive control of hyperbolic PDEs*. Springer, 2019.
- Jean Auriol and Florent Di Meglio. Robust output feedback stabilization for two heterodirectional linear coupled hyperbolic pdes. *Automatica*, 115:108896, 2020.
- Aw and Michel Rascle. Resurrection of “second order” models of traffic flow. *SIAM journal on applied mathematics*, 60(3):916–938, 2000.
- Francois Belletti, Mandy Huo, Xavier Litrico, and Alexandre M Bayen. Prediction of traffic convective instability with spectral analysis of the aw–rascle–zhang model. *Physics Letters A*, 379(38):2319–2330, 2015.
- Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Neural operators for bypassing gain and control computations in pde backstepping. *IEEE Transactions on Automatic Control*, 2023a.
- Luke Bhan, Yuanyuan Shi, and Miroslav Krstic. Operator learning for nonlinear adaptive control. In *Learning for Dynamics and Control Conference*, pages 346–357. PMLR, 2023b.
- Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE transactions on neural networks*, 6(4):911–917, 1995.
- André De Palma and Robin Lindsey. Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, 19(6):1377–1399, 2011.
- Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.
- Miroslav Krstic and Andrey Smyshlyaev. *Boundary control of PDEs: A course on backstepping designs*. SIAM, 2008.
- Miroslav Krstic, Luke Bhan, and Yuanyuan Shi. Neural operators of backstepping controller and observer gain functions for reaction–diffusion pdes. *Automatica*, 164:111649, 2024.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.
- Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deepnet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.
- Jie Qi, Jing Zhang, and Miroslav Krstic. Neural operators for pde backstepping control of first-order hyperbolic pde with recycle and delay. *Systems & Control Letters*, 185:105714, 2024.
- Martin Schönhof and Dirk Helbing. Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science*, 41(2):135–166, 2007.

- Yuanyuan Shi, Zongyi Li, Huan Yu, Drew Steeves, Anima Anandkumar, and Miroslav Krstic. Machine learning accelerated pde backstepping observers. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 5423–5428. IEEE, 2022.
- Rafael Vazquez, Miroslav Krstic, and Jean-Michel Coron. Backstepping boundary stabilization and state estimation of a 2×2 linear hyperbolic system. In *2011 50th IEEE conference on decision and control and european control conference*, pages 4937–4942. IEEE, 2011.
- Huan Yu and Miroslav Krstic. Traffic congestion control for aw–rasclé–zhang model. *Automatica*, 100:38–51, 2019.
- Huan Yu and Miroslav Krstic. *Traffic Congestion Control by PDE Backstepping*. Springer, 2022.
- H Michael Zhang. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, 36(3):275–290, 2002.
- Liguo Zhang, Christophe Prieur, and Junfei Qiao. Pi boundary control of linear hyperbolic balance laws with stabilization of arz traffic flow models. *Systems & Control Letters*, 123:85–91, 2019.