

Approximation and Generalization of DeepONets for Learning Operators Arising from a Class of Singularly Perturbed Problems

Ting Du¹, Zhongyi Huang¹ and Ye Li^{2,*}

¹Department of Mathematics Sciences, Tsinghua University, Beijing 100084, China.

²College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China.

Received 11 May 2023; Accepted (in revised version) 5 October 2023.

Abstract. Singularly perturbed problems present inherent difficulty due to the presence of thin layers in their solutions. To overcome this difficulty, we propose using deep operator networks (DeepONets), a method previously shown to be effective in approximating nonlinear operators between infinite-dimensional Banach spaces. In this paper, we demonstrate for the first time the application of DeepONets to one-dimensional singularly perturbed problems, achieving promising results that suggest their potential as a robust tool for solving this class of problems. We consider the convergence rate of the approximation error incurred by the operator networks in approximating the solution operator, and examine the generalization gap and empirical risk, all of which are shown to converge uniformly with respect to the perturbation parameter. By utilizing Shishkin mesh points as locations of the loss function, we conduct several numerical experiments that provide further support for the effectiveness of operator networks in capturing the singular layer behavior.

AMS subject classifications: 65M10, 78A48

Key words: Deep operator network, singularly perturbed problem, Shishkin mesh, uniform convergence.

1. Introduction

Singularly perturbed problems (SPP) are widely employed in diverse areas of applied mathematics, such as fluid dynamics, aerodynamics, meteorology, and modeling of semiconductor devices, among others. These problems incorporate a small parameter ε which typically appears before the highest order derivative term and reflects medium properties. As $\varepsilon \rightarrow 0$, the derivative (or higher-order derivative) of the solution to the problem becomes infinite, leading to the emergence of boundary or inner layers. In these regions, the

*Corresponding author. Email addresses: dt20@mails.tsinghua.edu.cn (T. Du), zhongyih@mail.tsinghua.edu.cn (Z. Huang), yeli20@nuaa.edu.cn (Y. Li)

solution (or its derivative) undergoes significant changes, while away from the layers, the behavior of the solution is regular and slow. Due to the presence of thin layers, conventional methods for solving this class of problems exhibit either high computational complexity or significant computational expense. For a comprehensive analysis and numerical treatment of singularly perturbed problems, refer to the books by Miller *et al.* [24] and Roos *et al.* [28].

Deep neural networks possess the universal approximation property, which allows them to approximate any continuous or measurable finite-dimensional function with arbitrary accuracy. Consequently, they have become a popular tool for solving partial differential equations (PDEs) by serving as a trial space. PDE solvers based on neural networks include Feynman-Kac formula-based methods [2, 6, 7] and physics-informed neural networks (PINNs) [27]. Moreover, machine learning techniques can be used for operator learning (i.e., mapping one infinite-dimensional Banach space to another) to approximate the underlying solution operator of an equation. To accomplish this, new network structures have been proposed, such as deep operator networks (DeepONets) [21] and neural operators [14, 16].

The broad adoption of machine learning techniques in solving PDEs has introduced novel approaches for handling singularly perturbed problems. However, the spectral bias phenomenon [26], which refers to the difficulty of neural networks in capturing sharp fluctuations of the function, poses a significant challenge for directly applying machine learning techniques to solve such problems. In [1], it was observed that the traditional PINNs approach did not produce a satisfactory solution nor capture the singular behavior of the boundary layer. To overcome this problem, a deformation of the traditional PINN based on singular perturbation theory is presented in [1].

Operator learning techniques have received little attention in the context of singularly perturbed problems, leaving an important research gap to be filled. Our focus is on DeepOnet, a simple yet powerful model. It originated from the universal approximation theorem proved by Chen *et al.* [4], where the shallow network was later extended to a deep network by Lu *et al.* [21], resulting in the proposal of DeepONets. This theorem was further extended to more general operators in [15], thereby providing theoretical assurance of the validity of DeepOnet to approximate the solution operator. Specifically, for any error bound ϵ , there exists an operator network that can approximate the operator with an accuracy not exceeding ϵ , while also guaranteeing that the size of the network is at most polynomially increasing concerning $1/\epsilon$. Theoretical and experimental results [5, 15] support the polynomially increasing property of DeepONets in some cases. DeepOnet has demonstrated promising performance in various applications, including fractional derivative operators [21], stochastic differential equations [21], and advection-diffusion equations [5]. To extend the applicability of DeepOnet to more general situations, several deformations have been proposed, including physics-informed DeepOnet [32], multiple-input DeepOnet [10], pre-trained DeepOnet for multi-physics systems [3, 23], proper orthogonal decomposition-based DeepOnet (POD-DeepOnet) [22], and Bayesian DeepOnet [17].

In this paper, we employ DeepONets to approximate the solution operator of the convection-diffusion problem, with particular attention to the case of Dirichlet boundary conditions. As it is known that for non-homogeneous Dirichlet boundary conditions, the solution