Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2021-0186

Vol. **31**, No. 4, pp. 1272-1295
April 2022

A Rate of Convergence of Physics Informed Neural Networks for the Linear Second Order Elliptic PDEs

Yuling Jiao[1,2], Yanming Lai[1], Dingwei Li[1], Xiliang Lu[1,2], Fengru Wang[1,2], Yang Wang[3] and Jerry Zhijian Yang[1,2,*]

[1] *School of Mathematics and Statistics, Wuhan University, P.R. China.*
[2] *Hubei Key Laboratory of Computational Science, Wuhan University, P.R. China.*
[3] *Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.*

Received 16 September 2021; Accepted (in revised version) 16 March 2022

Abstract. In recent years, physical informed neural networks (PINNs) have been shown to be a powerful tool for solving PDEs empirically. However, numerical analysis of PINNs is still missing. In this paper, we prove the convergence rate to PINNs for the second order elliptic equations with Dirichlet boundary condition, by establishing the upper bounds on the number of training samples, depth and width of the deep neural networks to achieve desired accuracy. The error of PINNs is decomposed into approximation error and statistical error, where the approximation error is given in C^2 norm with ReLU^3 networks (deep network with activation function $\max\{0,x^3\}$) and the statistical error is estimated by Rademacher complexity. We derive the bound on the Rademacher complexity of the non-Lipschitz composition of gradient norm with ReLU^3 network, which is of immense independent interest.

AMS subject classifications: 62G05, 65N12, 65N15, 68T07

Key words: PINNs, ReLU^3 neural network, B-splines, Rademacher complexity.

1 Introduction

Classical numerical methods such as the finite element method are successful to solve the low-dimensional PDEs, see e.g., [6, 7, 13, 24, 30]. However these methods may encounter some difficulties in both theoretical analysis and numerical implementation for

*Corresponding author. *Email addresses:* yulingjiaomath@whu.edu.cn (Y. Jiao), laiyanming@whu.edu.cn (Y. Lai), lidingv@whu.edu.cn (D. Li), xllv.math@whu.edu.cn (X. Lu), wangfr@whu.edu.cn (F. Wang), yangwang@ust.hk (Y. Wang), zjyang.math@whu.edu.cn (J. Z. Yang)

the high-dimensional PDEs. Motivated by the facts that deep learning method for high-dimensional data analysis has been achieved great successful applications in discriminative, generative and reinforcement learning [9,11,28], solving high dimensional PDEs with deep neural networks becomes an extremely potential approach and has attracted a lot of attentions [2,5,10,17,25,29,32,33]. Due to the excellent approximation ability of the deep neural networks, several numerical schemes have been proposed to solve PDEs with deep neural networks including the deep Ritz method (DRM) [32], physics-informed neural networks (PINNs) [25], and deep Galerkin method (DGM) [33]. Both DRM and DGM are applied to variational forms of PDEs, and PINNs are based on residual minimization to the differential equation, see [2,17,25,29], which can be extended to general PDEs [14,16,22,23].

Despite the above mentioned deep PDEs solvers work well empirically, rigorous numerical analysis for these methods are far from complete. The convergence rate of DRM with two layer networks and deep networks are studied in [8,12,19,20], the convergence of PINNs are given in [21,26,27]. In this work, we will provide the nonasymptotic convergence rate of the PINNs with ReLU^3 networks, i.e., a quantitative error estimation with respect to the topological structure of the neural networks (the depth and width) and the number of the samples. Hence it gives a rule to determine the hyper-parameters to achieve a desired accuracy. Our contributions are summarized as follows.

Our contributions and main results

- We obtain the approximation results ReLU^3 network in $C^2(\bar{\Omega})$, see Theorem 3.1, i.e., $\forall \bar{u}\in C^3(\bar{\Omega})$ and for any $\epsilon>0$, there exists a ReLU^3 network u_ϕ with depth $\lceil \log_2 d\rceil+2$ and width $C(d,\|\bar{u}\|_{C^3(\bar{\Omega})})\left(\frac{1}{\epsilon}\right)^d$ such that

$$\|\bar{u}-u_\phi\|_{C^2(\bar{\Omega})}\leq\epsilon,$$

  where d, $\bar{\Omega}$, $C(d,\|\bar{u}\|_{C^3(\hat{\Omega})})$ stands for the dimension of x, the closure of the domain Ω and some numerical constant that only depends on $(d,\|\bar{u}\|_{C^3(\hat{\Omega})})$, respectively.

- We establish an upper bound of the statistical error for PINNs by applying the tools of Pseudo dimension, especially we give an upper bound of the Rademacher complexity to the derivative of ReLU^3 network which is non-Lipschitz composition with ReLU^3 network, via calculating the Pseudo dimension of networks with ReLU, ReLU^2 and ReLU^3 activation functions, see Theorem 4.1. We prove that $\forall \mathcal{D},\mathcal{W}\in\mathbb{N}$ and $\epsilon>0$, if the number of training samples in PINNs is with the order $\mathcal{O}\left(\mathcal{D}^6\mathcal{W}^2(\mathcal{D}+\log\mathcal{W})\left(\frac{1}{\epsilon}\right)^{2+\delta}\right)$, where δ is an arbitrarily positive number, then the statistical error

$$\mathbb{E}_{\{X_k\}_{k=1}^N,\{Y_k\}_{k=1}^M}\sup_{u\in\mathcal{P}}\left|\mathcal{L}(u)-\widehat{\mathcal{L}}(u)\right|\leq\epsilon,$$

  where $\mathcal{L}$ and $\widehat{\mathcal{L}}$ are loss functions defined in (2.2) and (2.3) respectively.