Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2023-0063

Vol. 34, No. 3, pp. 813-836
September 2023

A Rate of Convergence of Weak Adversarial Neural Networks for the Second Order Parabolic PDEs

Yuling Jiao[1,2], Jerry Zhijian Yang[1,2,*], Cheng Yuan[1,2] and Junyu Zhou[1]

[1] *School of Mathematics and Statistics, Wuhan University, Wuhan, China.*
[2] *Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan, China.*

Received 4 March 2023; Accepted (in revised version) 2 June 2023

Abstract. In this paper, we give the first rigorous error estimation of the Weak Adversarial Neural Networks (WAN) in solving the second order parabolic PDEs. By decomposing the error into approximation error and statistical error, we first show the weak solution can be approximated by the $ReLU^2$ with arbitrary accuracy, then prove that the statistical error can also be efficiently bounded by the Rademacher complexity of the network functions, which can be further bounded by some integral related with the covering numbers and pseudo-dimension of $ReLU^2$ space. Finally, by combining the two bounds, we prove that the error of the WAN method can be well controlled if the depth and width of the neural network as well as the sample numbers have been properly selected. Our result also reveals some kind of freedom in choosing sample numbers on $\partial\Omega$ and in the time axis.

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

Key words: Weak Adversarial Networks, second order parabolic PDEs, error analysis.

1 Introduction

Partial differential equations (PDEs) have been widely and successfully applied to modeling in physics, chemistry and economics. Meanwhile, solving PDEs numerically has also drawn a large attention for practical simulation with PDEs. During the past decades, various numerical methods (e.g. the finite element method [1–5], discontinuous Galerkin method [6, 7], finite volume method [8, 9] and finite difference method [10, 11]) have been studied and used in solving low-dimensional problems. However, in the situation of high-dimensional PDEs, an issue known as the curse of dimensionality (CoD) would

*Corresponding author. *Email addresses:* yulingjiaomath@whu.edu.cn (Y. Jiao), zjyang.math@whu.edu.cn (J. Z. Yang), yuancheng@whu.edu.cn (C. Yuan), junyuzh@whu.edu.cn (J. Zhou)

make the traditional methods impractical, in the sense that the number of freedom to achieve a given precision in solving PDEs may increase exponentially with the dimension.

On the other hand, with recent development of machine learning in high-dimensional data analysis, several works based on artificial neural networks have been proposed to solve high-dimensional partial differential equations, including the Deep Ritz Method (DRM) [12], the Physics-Informed Neural Networks (PINNs) [13] and the Weak Adversarial Neural Networks (WAN) [14]. Both DRM and WAN use the variational forms of the PDEs, while the PINNs is based on the square residual of the differential equations, see [13, 15–17]. Further generalizations of PINNs have been explored in [18–21].

In addition to the successful practical application and good performance, several works on error estimation for DRM structure [22–25] and PINNs structure [26–30] have also been conducted, while a rigorous numerical analysis for WAN is still needed. Due to the weak formulation, the WAN structure has more benefits since it doesn't need a strong solution of the problem, which can allow us to choose the test function φ to make the process more accurate. In this paper, we would give the first systematic error analysis for the WAN structure in solving the second order parabolic PDEs. We decompose the error into approximation error and statistical error. For the approximation error, we show that the weak solution can be arbitrarily approximated by the $ReLU^2$ network functions with sufficiently large depth and width. To bound the statistical error, we introduce the Rademacher complexity of the non-Lipschitz composition of derivatives, of which the upper bound can be further obtained via using the concept of covering number and pseudo-dimension.

In summary, our main contributions are as follows:

- Based on the classical Babuska-Brezzi theory [31] for parabolic problems, we bound the error of WAN by the risk of optimal empirical solution, then we decompose the error in risk into approximation error and statistical error, to which some recent neural network approximation theory and learning theory can be applied.

- We obtain the upper bounds of the statistical error by using the concept of Rademacher complexity, covering number and pseudo-dimension [32, 33]. The main difficulty here is to bound the non-Lipschitz composition of the derivative. The result tells us how to get the desired accuracy. See Theorem 5.1. Let d be the dimension of the problem, $\mathcal{D},\mathcal{W}\in\mathbb{N}$ be the depth and width of the neural network function class we chose to approximate the weak solution and the test function, $\mathcal{B}\in\mathbb{R}^+$ is constant. For any $\varepsilon\geq 0$, if the number of sample satisfying:

$$\begin{cases} N=C(d,|\Omega|,\mathcal{B})\mathcal{D}^2\mathcal{W}^2(\mathcal{D}+\log\mathcal{W})\left(\frac{1}{\varepsilon}\right)^{2+\delta}, \\ M=C(d,|\partial\Omega|,\mathcal{B})f_M(\mathcal{D},\mathcal{W})\left(\frac{1}{\varepsilon}\right)^{k_1}, \\ L=C(d,|T|,\mathcal{B})\mathcal{D}^2 f_L(\mathcal{D},\mathcal{W})\left(\frac{1}{\varepsilon}\right)^{k_2}, \end{cases}$$