Ann. Appl. Math.
doi: 10.4208/aam.OA-2025-0028

Vol. **41**, No. 3, pp. 313-376
August 2025

On Residual Minimization for PDEs: Failure of PINN, Modified Equation, and Implicit Bias

Tao Luo[1,2], Qixuan Zhou [1,*]

[1] *School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, 200240, China.*

[2] *Institute of Natural Sciences, CMA-Shanghai, MOE-LSC and Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai, 200240, China*

Received 12 August 2025

Abstract. As a popular and easy-to-implement machine learning method for solving differential equations, the physics-informed neural network (PINN) sometimes may fail and find poor solutions which bias against the exact ones. In this paper, we establish a framework of modified equation to explain the failure phenomenon and characterize the implicit bias of a general residual minimization (RM) method. We provide a simple way to derive the modified equation which models the numerical solution obtained by RM methods. Next, we show the modified solution deviates from the original exact solution. The proof uses a by-product of this paper, that is, a necessary and sufficient condition on characterizing the singularity of the coefficients. This equivalent condition can be extended to other types of equations in the future. Finally, we prove, as a complete characterization of the implicit bias, that RM method implicitly biases the numerical solution against the exact solution and towards a modified solution. In this work, we focus on elliptic equations with discontinuous coefficients, but our approach can be extended to other types of equations and our understanding of the implicit bias may shed light on further development of deep learning based methods for solving equations.

AMS subject classifications: 35D30, 35D35, 35R05, 35R06, 65N15

Key words: residual minimization, PINN, implicit bias, modified equation.

---

*Corresponding author.
Emails: luotao41@sjtu.edu.cn (T. Luo), zhouqixuan@sjtu.edu.cn(Q. Zhou)

1 Introduction

The application of machine learning, particularly deep neural networks (DNNs), has gained significant attention in recent years for solving partial differential equations (PDEs) [1–7]. Compared with traditional numerical schemes, such as finite difference, finite element methods, and spectral methods, which are often limited by the "curse of dimensionality," DNNs have demonstrated success in solving many high-dimensional problems [8]. Although the traditional numerical methods are powerful for low-dimensional problems, it can be challenging to design a proper scheme to solve low-dimensional problems with low-regularity solutions or boundaries [9–14]. Therefore, DNNs are also promising in solving low-dimensional problems with low-regularity solutions or complex boundaries, such as problems with discontinuous elastic or dielectric constants in composite materials.

A widely used approach to solving PDEs is to utilize DNNs to parameterize the solution and optimize the parameters in an objective function, which is usually formulated as a least-squares or variational loss function (also known as a risk function). The physics-informed neural network (PINN) method was first proposed in the 1990s [15], later studied by Sirignano and Spiliopoulos under the name Deep Galerkin Method (DGM) [16], and popularized as PINN by Raissi et al. [6]. In this method, a DNN is trained to minimize the sum of the residuals of the PDE and the boundary condition. The Deep Ritz Method (DRM) [1] instead adopts a variational formulation, minimizing an energy functional. Many other approaches have also been proposed for solving PDEs with neural networks [17–19]. For further advances in PINN, we refer readers to the review articles [20, 21] and the references therein.

For completeness, we also mention that operator learning has shown promise in solving both forward and inverse PDE problems [22–25]. Despite this growing diversity, PINN has received particular attention due to its simplicity and ease of implementation: its risk function is merely the residual of the PDE and boundary conditions, without requiring the variational form needed in DRM, which is often unavailable in practical applications.

A theoretical study of DNNs is crucial for understanding and improving PDE solvers based on neural networks. For instance, the universal approximation theorem [26] guarantees the ability of wide networks to approximate continuous functions. Subsequent works have extended this by quantifying approximation rates [49], and analyzing PINN approximation errors in the case of smooth elliptic problems [27]. Several studies have derived generalization bounds under strong regularity assumptions [28–31], supporting convergence results when the solution is smooth and the network is sufficiently expressive.

However, many real-world PDEs involve solutions of low regularity, due to discontinuities in coefficients or domain geometry, etc. Several recent studies have investigated