Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2024-0153

Vol. **39**, No. 3, pp. 696-712
March 2026

Weak Minimization-Sampling Network for Solving PDEs

Zhenqian E[1,*] and Weiguo Gao[2,3]

[1] *School of Data Science, Fudan University, Shanghai 200000, P.R. China.*
[2] *School of Mathematical Sciences, Fudan University, Shanghai 200000, P.R. China.*
[3] *Shanghai Key Laboratory for Contemporary Applied Mathematics, Fudan University, Shanghai 200433, P.R. China.*

Received 11 July 2024; Accepted (in revised version) 2 August 2024

Abstract. In many engineering applications, people concentrate on deriving the weak solution to a given partial differential equation (PDE) model. Traditionally, both the trial space and the test space are constructed as the span of a set of basis functions with compact supports and the approximate solution is achieved by solving a (linear) system of equations. A recent advancement in this field is the Weak Adversarial Network (WAN), a machine learning-based method that employs deep neural networks (DNNs) to approximate trial and test spaces. WAN addresses PDEs through a min-max optimization framework, showcasing its effectiveness in solving high-dimensional PDEs and surpassing the limitations of traditional numerical methods. In this paper, motivated by WAN, we propose a Weak Minimization-Sampling Network (WMSN) for solving PDEs which also leverages DNNs to approximate the trial and test spaces. Instead of the min-max loss function, our key idea is to reformulate the weak formulation into a loss function defined as an expectation over all possible DNNs in the test space with only the network parameters in the trial space to be trained. This leads to a minimization problem in the trial space and a sampling problem over the test space, which can be solved by stochastic gradient descent type methods with neat implementation. Through a series of numerical tests, we find that WMSN exhibits comparable or slightly better results with smaller computational cost when compared to WAN.

AMS subject classifications: 35D30, 35J25, 35J66, 68T07
Key words: Partial differential equations, weak formulation, deep neural network, random sampling.

1 Introduction

Solving high-dimensional partial differential equations has been a longstanding challenge in scientific computing. Classical numerical methods, including the finite differ-

*Corresponding author. *Email addresses:* zqe18@fudan.edu.cn (Z. E), wggao@fudan.edu.cn (W. Gao)

ence method [1], finite element method [2] and spectral method [3], generally convert the PDE model into its discrete counterparts using grids or meshes in the computational domain. Subsequently, they solve the resulting systems of equations to determine the state variables at grid nodes or coefficients of functional expansions representing these variables. These methods have proven successful in many scenarios over the past decades, however, are hindered by the "curse of dimensionality" which manifests as an exponential growth in degrees of freedom with increasing dimensionality. Moreover, mesh generation is time-consuming. Therefore, traditional numerical methods only work well in low-dimensional settings ($d \leq 3$).

Recent advancements in deep learning have led to significant breakthroughs in various artificial intelligence tasks. Underlying this success is the potential ability of deep neural networks to approximate complex functions using a composition of linear transformations and non-linear activation functions, especially in high-dimensional spaces. Importantly, the growth rate of the number of parameters in these networks remains at most polynomial with respect to the dimensionality of the input space. This scalability proves particularly advantageous in contexts where constructing high-dimensional functions is essential, such as in the solution of high-dimensional partial differential equations. As a result, numerous DNN-based methodologies [4–23, 25, 26] have emerged in this area, showing promising avenues for further exploration and applications.

Here, we highlight several methods from the literature, categorized into three main classes. One prominent DNN-based approach relies on the variational formulation of PDEs to set up the learning problem. A notable numerical technique within this domain is the Ritz method. Inspired by this foundation, [4] proposes the deep Ritz method (DRM) which uses the variational formulation and the integral of the boundary residual as the loss function to optimize the parameters of the trial neural network solution. Deep Nitsche method [5] incorporates the idea of Nitsche into the framework of DRM to handle essential boundary conditions and provides the error estimate in the energy norm. Deep domain decomposition method [6] parallelizes the computation of DRM, thereby achieving enhanced accuracy while reducing computational costs. ALDL [7] rewrites the original problem as a minimax problem associated with a feasible augmented Lagrangian and then expresses the primal and dual variables with two individual DNNs respectively whose parameters are trained using the stochastic optimization method together with a projection technique. In [8], a variational neural network approach is proposed to address the Non-Newtonian ice flow model, incorporating a re-normalization technique and adaptive boundary penalties. However, it's important to note that variational formulations do not universally guarantee existence.

Another category involves designing algorithms that adhere to the strong form and aim to minimize the residual of the original partial differential equation. Deep Galerkin method (DGM) [9] formulates the squared residuals in the least-squares sense as the primary loss function. Additionally, boundary conditions are incorporated into the optimization process as penalty terms. Physical-informed neural networks (PINNs) [10] approximates the PDE solutions by integrating observed data points, differential equations,