Journal of Computational Mathematics
Vol.44, No.1, 2026, 1–34.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2407-m2024-0047

ELEMENT LEARNING: A SYSTEMATIC APPROACH OF ACCELERATING FINITE ELEMENT-TYPE METHODS VIA MACHINE LEARNING, WITH APPLICATIONS TO RADIATIVE TRANSFER*

Shukai Du[1)]
Department of Mathematics, University of Wisconsin–Madison, USA
Email: sdu49@wisc.edu
Samuel N. Stechmann
Department of Mathematics, Department of Atmospheric and Oceanic Sciences,
University of Wisconsin–Madison, USA

Abstract

In this paper, we propose a systematic approach for accelerating finite element-type methods by machine learning for the numerical solution of partial differential equations (PDEs). The main idea is to use a neural network to learn the solution map of the PDEs and to do so in an element-wise fashion. This map takes input of the element geometry and the PDE's parameters on that element, and gives output of two operators: (1) the in2out operator for inter-element communication, and (2) the in2sol operator (Green's function) for element-wise solution recovery. A significant advantage of this approach is that, once trained, this network can be used for the numerical solution of the PDE for any domain geometry and any parameter distribution without retraining. Also, the training is significantly simpler since it is done on the element level instead on the entire domain. We call this approach element learning. This method is closely related to hybridizable discontinuous Galerkin (HDG) methods in the sense that the local solvers of HDG are replaced by machine learning approaches. Numerical tests are presented for an example PDE, the radiative transfer or radiation transport equation, in a variety of scenarios with idealized or realistic cloud fields, with smooth or sharp gradient in the cloud boundary transition. Under a fixed accuracy level of 10^{-3} in the relative L^2 error, and polynomial degree $p = 6$ in each element, we observe an approximately 5 to 10 times speed-up by element learning compared to a classical finite element-type method.

Mathematics subject classification: 65N30, 65N55, 68T07.
Key words: Scientific machine learning, Spectral element, Discontinuous Galerkin, Hybridization, Hybridizable discontinuous Galerkin, Radiation transport, Radiative transfer.

1. Introduction

1.1. Background and motivation

In the past decade, (artificial) neural networks and machine learning tools have surfaced as game-changing technologies across numerous fields, resolving an array of challenging problems.

* Received March 5, 2024 / Revised version received June 1, 2024 / Accepted July 15, 2024 / Published online October 8, 2024 /
1) Corresponding author

Examples include image recognition [61,66], playing the game go [97], protein folding [56], and large language models such as GPT3 [9].

Given these impressive results, it is reasonable to envision the potential of neural networks (NNs) for the numerical solution of partial differential equations or other scientific computing problems. There has been significant work in this direction [6, 10, 11, 15–17, 25, 27, 29, 40, 41, 44, 45, 48, 53, 55, 57–59, 71, 81, 83, 86, 94, 100, 103, 106, 108, 109]. As examples, we mention two major groups – (1) neural networks as function approximators, (2) neural networks as operator approximators.

The first type of these methods uses neural networks to approximate the solution of PDEs. Examples include physics-informed neural networks (PINNs) [24,73,75,87] and Deep Ritz methods [107]. These methods have demonstrated promising results, especially in the realm of high-dimensional problems or inverse problems, since they seem to bypass the notorious "curse of dimensionality", potentially attributable to the neural network's efficient approximation capabilities in high-dimensional spaces [36, 47, 78].

While the potential advantages are promising, it remains uncertain if the aforementioned methods have a real advantage over traditional algorithms, such as finite element-type methods, when addressing many classic PDEs. A crucial factor contributing to this uncertainty stems from the complex optimization problems and error landscapes-often non-convex-introduced by neural networks due to their hidden layers and non-linear activation functions.

The second type of these methods employs neural networks to approximate operators, with methods like Neural Operator [60,76] and DeepOnet [72] serving as leading examples. A primary advantage of these methods hinges on their speed – once trained, solving a problem is simply a forward propagation of the network. For instance, in [82], it was reported that the neural operator based method can be orders of magnitude faster than classical finite volume/difference based methods in weather simulation. This suggests great potential of using machine learning to accelerate computation.

However, this type of approach is essentially data-driven and its performance can depend on the training samples and how well it generalizes. Consequently, the reliability of the result may falter during extreme events, as indicated in [82]. In addition to this, it is common for these approaches to be constrained by the geometry of the domain and boundary conditions. For instance, the Fourier neural operator method that relies on Fourier expansion for solution representation is specifically tailored for rectangular/cubic domains with periodic boundary conditions [60]. However, there are recent attempts to extend operator learning to more complex geometries [70, 95].

On the other hand, for many years, traditional methods like finite element-type methods have played a critical role in the advancement of various scientific and engineering disciplines. These methodologies, honed and well-understood over more than five decades, have given us a plethora of reliable and robust techniques successfully employed across an array of problems. Examples include, but are definitely not limited to, conforming and non-conforming finite element [8, 23], mixed finite element [80, 89], discontinuous Galerkin [2, 21], along with effective techniques such as slope limiter [21, 69], multi-scale finite elements [38], finite element exterior calculus [3, 43], *hp*-adaptivity [5, 54] to tailor these methods to different application scenarios [22, 33–35, 42, 50, 74, 77, 101].

Despite the undeniable utility of these classical approaches, they also come with constraints. A primary limitation is on their speed, especially when facing high-dimensional problems, such as those found in radiative transfer, or scenarios when fast forward solvers are necessary, such