INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 19, Number 5, Pages 603–629

COMPARATIVE STUDIES ON MESH-FREE DEEP NEURAL NETWORK APPROACH VERSUS FINITE ELEMENT METHOD FOR SOLVING COUPLED NONLINEAR HYPERBOLIC/WAVE EQUATIONS

XINGWEN ZHU, MINGYAN HE, AND PENGTAO SUN*

Abstract. In this paper, both the finite element method (FEM) and the mesh-free deep neural network (DNN) approach are studied in a comparative fashion for solving two types of coupled nonlinear hyperbolic/wave partial differential equations (PDEs) in a space of high dimension $\mathbb{R}^d$ $(d > 1)$, where the first PDE system to be studied is the coupled nonlinear Korteweg-De Vries (KdV) equations modeling the solitary wave and waves on shallow water surfaces, and the second PDE system is the coupled nonlinear Klein-Gordon (KG) equations modeling solitons as well as solitary waves. A fully connected, feedforward, multi-layer, mesh-free DNN approach is developed for both coupled nonlinear PDEs by reformulating each PDE model as a least-squares (LS) problem based upon DNN-approximated solutions and then optimizing the LS problem using a $(d+1)$-dimensional space-time sample point (training) set. Mathematically, both coupled nonlinear hyperbolic problems own significant differences in their respective PDE theories; numerically, they are approximated by virtue of a fully connected, feedforward DNN structure in a uniform fashion. As a contrast, a distinct and sophisticated FEM is developed for each coupled nonlinear hyperbolic system, respectively, by means of the Galerkin approximation in space and the finite difference scheme in time to account for different characteristics of each hyperbolic PDE system. Overall, comparing with the subtly developed, problem-dependent FEM, the proposed mesh-free DNN method can be uniformly developed for both coupled nonlinear hyperbolic systems with ease and without a need of mesh generation, though, the FEM can produce a concrete convergence order with respect to the mesh size and the time step size, and can even preserve the total energy for KG equations, whereas the DNN approach cannot show a definite convergence pattern in terms of parameters of the adopted DNN structure but only a universal approximation property indicated by a relatively small error that rarely changes in magnitude, let alone the dissipation of DNN-approximated energy for KG equations. Both approaches have their respective pros and cons, which are also validated in numerical experiments by comparing convergent accuracies of the developed FEMs and approximation performances of the proposed mesh-free DNN method for both hyperbolic/wave equations based upon different types of discretization parameters changing in doubling, and specifically, comparing discrete energies obtained from both approaches for KG equations.

Key words. Coupled hyperbolic/wave equations, Korteweg-De Vries (KdV) equations, Klein-Gordon (KG) equations, deep neural network (DNN), finite element method (FEM), space-time sample points (training) set, least-squares (LS), convergence accuracy, energy conservation.

1. Introduction

In this paper, we choose the following two types of coupled nonlinear hyperbolic/wave partial differential equations (PDEs) as two model problems to be studied, numerically: Korteweg-De Vries (KdV) equations and Klein-Gordon (KG) equations, both are coupled hyperbolic system defined in a space of high dimension $\mathbb{R}^d$ $(d > 1)$ and in a one-dimensional time interval $[0, T]$. The KdV equation was

Received by the editors September 16, 2021 and, in revised form, March 24, 2022; accepted May 27, 2022.
2000 *Mathematics Subject Classification.* 65M60, 65N30, 35L57, 35L75, 35L70, 68Q32, 35Q93.
*Corresponding author.

first introduced in 1985 by Korteweg and de Vries for a (1+1)-dimensional case, later the (d+1)-dimensional cases of coupled KdV equations were developed to explain more involved nonlinear phenomena [24, 39, 40, 46, 69, 51]. The KdV equation(s) is a very important hyperbolic PDE model, both mathematically and practically, for the description of small amplitude shallow-water waves with weakly nonlinear restoring forces, long internal waves in a density-stratified ocean, ion acoustic waves in a plasma, solitary waves on the intensity of light in optical fibers, acoustic waves on a crystal lattice, and fluctuation phenomena in biological and physical systems [9, 35, 64]. As for the second hyperbolic problem to be studied in this paper, KG equations, plays a significant role in many scientific applications as well, such as in studying solitons and condensed matter physics [6, 55, 8, 4, 1], in investigating the interaction of solitons in a collisionless plasma [31, 42], and in examining the recurrence of initial states and the nonlinear wave equations [15, 62, 61, 60, 45].

Both KdV and KG equations are targeted together in this paper because not only their numerical studies are of great scientific significance and research interests but also: (1) both of them fall into the same category of nonlinear hyperbolic/wave problems, in general; (2) each one of them owns respective special properties in PDE theories, in particular; (3) a stable, convergent and/or energy-preserving finite element method (FEM) for each PDE model is uneasy and subtle to be developed. While a comprehensive numerical analysis for each one of these two PDE systems is still a hot topic even now and somehow challenging [3, 67, 19, 33, 30, 34, 23, 11, 28], in this paper we are dedicated to developing FEMs and recently emerging mesh-free, deep neural network (DNN) approaches for both coupled nonlinear hyperbolic/wave problems, respectively, to comparing numerical complexities and convergence properties of both FEM and DNN approach, and finally, to reaching a comparative conclusion based on these two numerical approaches for two distinct problems that belong to the same kind of PDEs, largely.

DNN has been demonstrated as a powerful tool to conquer the curse of dimensionality [16, 18, 25, 63], and have been applied to solve PDEs, e.g., the deep BSDE method [20, 27], the deep Galerkin method (DGM) [59], the physics-informed neural networks (PINN) [48, 44, 49, 32], the deep Ritz method (DRM) [21], the weak adversarial networks (WAN) [68], and the deep Nitsche method (DNM) [38]. The deep BSDE reformulates the time-dependent equations into stochastic optimization problems. The DGM and the PINN train neural networks by minimizing the mean squared error loss of the PDE residual, while the DRM trains networks by minimizing the energy functional of the variational problem that is equivalent to the targeted PDE. The WAN adopts the weak form of original PDEs and trains the primary and adversarial network alternatively with the min-max weak form, and, the DNM enhances the DRM with natural treatment of essential boundary conditions. Recently, an additional neural network is trained to impose Dirichlet boundary conditions [57].

In this paper, we employ a likelihood of the DGM and/or of the PINN approaches to solve both coupled nonlinear hyperbolic systems by adopting DNN functionals to approximate unknown variables, reformulating each PDE model and its initial & boundary conditions as a series of least-squares (LS) problems in the mean squared error form, whose summation defines a total loss function, then minimizing this loss function with a standard optimization algorithms such as the stochastic gradient descent (SGD) method [53, 7, 43, 37, 54], the trust region method [13, 52, 65] or the derivative-free method [50, 2, 66] based upon a $(d+1)$-dimensional space-time