Journal of Computational Science &Engineering 61 (2023) 23-31

Available at http://www.asocse.org ISSN 1710-4068

A PINN-based method for solving the convective diffusion equation

Rujia Liu1, Yin Huang[1], Jieyu Ding [2,3*]

[1]College of Computer Science and Technology, Qingdao University, Qingdao 266071, China

[2]School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

[3]Center for Computational Mechanics and Engineering Simulation, Qingdao University, Qingdao 266071, China

Abstract: **Convective diffusion equations are a special class of partial differential equations, and the study of their numerical solution methods has been a hot issue. PINNs (Physics-informed Neural Networks) based on physical information are introduced and applied to solve two-dimensional convective diffusion equations. In contrast to traditional neural network models, PINNs introduce automatic differentiation techniques in the model building process, and compile the physical information, i.e., partial differential equation information, to obtain the optimization objectives of the neural network parameters, i.e., weights and biases, in the model by defining loss functions, which are then solved using existing optimization**

Keywords: Partial differential equations； fully connected neural networks; convective diffusion equations; Dirichlet boundaries; PINN methods

Received 3 May 2023, Revised 7 May 2023, Accepted 28 May 2023

*Corresponding Author: Jieyu Ding

1. Introduction

Convective diffusion equations are a key branch of partial differential equations and are widely used in many fields, such as environmental science and fluid mechanics, to describe heat transfer, material diffusion, and dispersion of liquid solutes in drainage membranes, making it particularly important to study numerical computational methods for convective diffusion problems [1- 3]. A large number of finite difference solution methods and finite element solution methods and theoretical analyses have been used to solve convective diffusion equations [4-8].

There are two very important parts in the convective diffusion equation: one is the diffusion term and the other is the convection term. The diffusion term describes the fluid migration or substance conduction due to the inhomogeneous distribution of a physical quantity in the fluid. When a physical quantity in a fluid is spatially unevenly distributed, diffusion leads to the diffusion of matter from regions of high concentration to regions of low concentration [9-10] in order to homogenize the distribution of the physical quantity, with the result that the spatial distribution of the physical quantity is altered. The convective term, on the other hand, describes how the macroscopic motion of the fluid leads to a change in the spatial distribution of physical quantities in it. When the fluid undergoes macroscopic motion, the relative positions between the different components change, and this change in relative position in turn affects the distribution of physical quantities in the fluid. Therefore, by analyzing some features in the convective diffusion equation one can derive the laws of variation of some physical quantities (such as concentration and energy) corresponding to the object [16]

In recent years, with the rapid development of machine learning algorithms, a class of PINNs [11] (Physics-Informed NeuralNetworks), a class of physics-informed neural network-based algorithms, has been proposed and widely used in various fields [12-14]. The most commonly used numerical computational method for solving hydrodynamic equations is the finite difference method, in which the basic difference formats are the windward format, Lax-Friedrichs format, Lax-Wendroff format, QUICK format, exponential difference format, and Crank-Nicolson format [19]. - Nicolson format [19]. These calculation methods are simple in format and low in

2023@ The American Computational Science Society. All rights reserved.

computational accuracy, and the results are not ideal for some complex problems and cannot meet the needs of actual numerical simulations. Later, Tian et al [20] proposed a high-precision exponential difference method, and Ding and Zhang constructed a spatial fourth-order and temporal fifth-order difference format, which uses semi-discrete and Padé approximation in time and polynomial tight difference in space as in the literature [21]. In [22], a high-precision fully implicit tight difference format is proposed for the two-dimensional non-constant incompressible vortex-velocity Navier-Stokes system of equations. In [23], a second-order five-point difference format is proposed for the two-dimensional convective diffusion equation, and in [24], a fourth-order accurate implicit compact difference format is obtained for solving the two-dimensional convective diffusion equation using the fourth-order Padé-type compact difference approximation of the first- and second-order derivatives, combined with the original equation. In [25], a high-precision compact difference format for solving the one-dimensional convective diffusion response problem on a nonuniform grid was constructed based on the Taylor series expansion of the function, in [26], a higher-order compact difference format for the one-dimensional convective diffusion equation was derived based on the Taylor series expansion of the function on a nonuniform grid, and in [27], a three-point fourth-order compact difference format on a nonuniform grid was proposed.

However, unlike traditional methods, PINNs define a loss function with respect to the root mean square error, i.e., the objective function, in order to combine the boundary conditions and the initial conditions. By optimizing this loss function, we can find the optimal parameters of the neural network, including weights and biases. By optimally solving for the optimal parameters, we also obtain a deep neural network model that can approximate the description of this physical-mechanical system. The core idea of this approach is to transform the physical-mechanical behavior into a learning problem for the neural network and to train the network to fit and approximate the behavior of the physical system. By optimizing the loss function, we can obtain optimal parameters that allow the neural network model to approximate the dynamics of the physical system to some extent. This approach, which integrates deep learning and physical modeling, has a wide range of applications. It can help us better understand and predict the behavior of physical systems, and also provides a new way to address the complexity and computational efficiency problems encountered in traditional physical modeling. By combining neural networks with physical equations, we can apply this approach to a variety of practical problems in practice, thus advancing scientific research and engineering applications.

This method greatly reduces the need for the amount of data by introducing additional physical information and shows better robustness for a deterministic model [15].

2. Two-dimensional convective diffusion equation

2.1 Convection equation

In practical applications, convective equations are often used together with other equations and conditions to describe specific physical processes. The general convective equation can be expressed as

$$\frac{\partial \phi}{\partial t} + v \nabla \phi = 0 \qquad (1)$$

where ϕ is the physical quantity to be found, such as the temperature or concentration of the fluid, t is the time, v is the velocity vector, and ∇ is the gradient operator.

The convection equation describes the variation of the physical quantity ϕ with time and space. By solving the convection equation, we can study the propagation and distribution of physical quantities in the fluid and the relationship with time.

2.2 Diffusion equation

In practical applications, diffusion equations are often used to describe the transport and diffusion phenomena of

2023@ The American Computational Science Society. All rights reserved.

substances, such as the diffusion of substances in chemical reactions and the temperature diffusion in heat conduction. By adjusting the diffusion coefficients and initial conditions in the equation, we can study the diffusion behavior under different conditions, optimize the efficiency of the diffusion process, or solve related engineering and scientific problems.

The general diffusion equation can be expressed as

$$\frac{\partial \phi}{\partial t} = D\nabla^2\phi \qquad (2)$$

where ϕ is the physical quantity to be found, such as concentration or temperature, t is time, D is the diffusion coefficient, and ∇^2 is the Laplace operator. The diffusion process is determined by the diffusion coefficient D, which describes the rate and intensity of the substance diffusion.

2.3 Convective diffusion equation

When considering the inclusion of Dirichlet boundary conditions, the equations describe how a fluid or substance propagates and diffuses in a two-dimensional region for known values on a given boundary.

The two-dimensional convective diffusion equation in general form containing the Dirichlet boundary can be expressed a

$$\frac{\partial \mathrm{u}}{\partial \mathrm{t}} = \mathrm{D}\left(\frac{\partial^2 \mathrm{u}}{\partial \mathrm{x}^2} + \frac{\partial^2 \mathrm{u}}{\partial \mathrm{y}^2}\right) - \mathrm{V}\left(\frac{\partial \mathrm{u}}{\partial \mathrm{x}} + \frac{\partial \mathrm{u}}{\partial \mathrm{y}}\right) + \mathrm{F}(\mathrm{x},\mathrm{y},\mathrm{t}) \qquad (3)$$

In this equation, u is the variable to be found, which can represent the concentration, temperature, or other properties of the fluid or substance. t is time, and x and y are spatial coordinates. d is the diffusion coefficient, which describes the intensity of the substance diffusion. v is the convective velocity, which describes the speed and direction of motion of the fluid or substance. $f(x, y, t)$ is the source term, which represents an external source or sink at a given location and time.

3. Neural networks based on physical information

In this paper, we propose the Physical Information Neural Network (PINN) method. PINN is able to combine the physical equations with a neural network model of the loss function, so that the optimization process is constrained by the physical equations, thus training a model that can satisfy the given physical laws. The main purpose of this method is to approximate the solution of partial differential equations.

3.1 PINN Module

Mathematically understood, a deep neural network is a structure composed of special functions. The simplest of these neural networks is the feedforward neural network (FNN), also known as the multilayer perceptron (MLP). Although many other types of neural networks, such as convolutional neural networks and recurrent neural networks, have been developed in the past decades, fully connected neural network remains one of the simplest and fastest solutions for most partial differential equations. Therefore, in this paper, the standard fully connected neural networks are used to solve physically constrained neural networks (PINNs) for partial differential equations. As an example, a standard fully connected neural network for the two-dimensional convection-diffusion equation is shown in Figure 1.

2023@ The American Computational Science Society. All rights reserved.

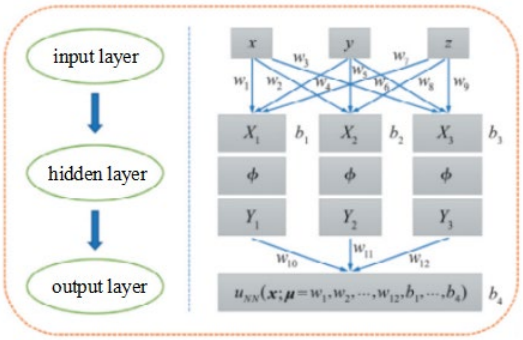

Figure 1.A simple neural network with a single hidden layer and three neurons

Unlike traditional neural network algorithms, in PINNs the loss function is defined as follows, i.e.

$$LOSS = LOSS_{data} + LOSS_F + LOSS_{boundary} \quad (4)$$

Where $LOSS_{data}$ corresponds to the loss function for the given data, $LOSS_F$ corresponds to the residual loss function, and $LOSS_{boundary}$ corresponds to the boundary loss function.

In the process of neural network training, the desired effect is achieved by minimizing the loss function. In this paper, we add weight coefficients to the loss function, and through experiments, we find that increasing the weight coefficients and appropriately decreasing the learning rate can make the neural network training results better and the absolute error becomes smaller.

4. Experiment

4.1 Data set and evaluation metrics

In order to test the practical applicability of our proposed method, we can apply it to the solution of two-dimensional partial differential equations, such as the seepage drive problem. Oil recovery has attracted widespread social attention, and the problem of seepage drive is currently of interest to many countries and regions around the world, and the problem has become an important issue in the international community. In fact, when water is injected into an oil formation, a movement of water-oil mixture is formed, and this movement is called seepage drive [18], therefore, the solution of this problem is of great importance for oil recovery. However, it is not an easy task to solve the seepage drive problem, which needs to be calculated and analyzed by applying the convective diffusion equation, and therefore, the seepage drive problem has become an important topic in the field of scientific research. This problem has attracted great attention in many countries and regions, therefore, we need to adopt an effective method to solve this problem, and our proposed method can provide a new way of thinking and solution for researchers. Therefore, our research results have an important theoretical and practical value. [17]

We consider the convective diffusion equations of the following form for the computational region $\Omega \in [0,1]$ containing the Dirichlet boundary, $g(x,y)$ and $F(x,y)$ given by the analytical solutions.

$$\begin{cases} Lu = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + pu_x + qu_y = F(x,y), & in\ the\Omega \\ u(x,y) = g(x,y), on\ the\ \partial\Omega \end{cases} \quad (5)$$

The analytical solution of the equation is

$$u(x,y) = e^{-a((x-0.5)^2+(y-0.5)^2)} + 0.5 \quad (6)$$

Among them, the meaning of each parameter is explained as shown in the table:

Table 1: Meaning of the parameters of the mathematical formula

Symbols	Description
u	Distribution of physical quantities (e.g., concentration, temperature, etc.) in two dimensions
x	Denotes the spatial coordinates, which correspond to the position of the horizontal axis in

2023@ The American Computational Science Society. All rights reserved.

	the two-dimensional case, respectively
y	Denotes the spatial coordinates, which correspond to the position of the vertical axis in the two-dimensional case, respectively
a	Diffusion coefficient, describing the degree of diffusion of a physical quantity in space
p	The component of convective velocity in the 'x' direction, describing the flow properties of the fluid in space
q	The component of convective velocity in the 'y' direction, describing the flow properties of the fluid in space

In the experimental validation phase, $p = 2$, $q = 2$, and $a = 5$ were taken to determine the steepness of the two-dimensional convective diffusion equation, and Mean Square Error (MSE) was used as an evaluation indicator plot, which is the mean value of the square of the difference between the predicted and true values for each sample. In statistics, MSE is an important indicator of the precision of an estimator. If we use an estimator to predict the value of an unknown parameter or variable, the MSE can tell us the magnitude of the bias and variance of this estimator. Usually, we would like the estimator to have both a small bias and a small variance, which makes the estimation result more accurate and reliable. The MSE is very effective for the performance of the learning method of physical mechanics system, which is calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$

where y_i denotes the predicted value or the value of the model output, $\hat{y}_i$ denotes the true or target value, and n denotes the number of samples.

The smaller the value of MSE, the smaller the difference between the predicted and true values, and the better the predictive performance of the model.MSE is sensitive to outliers because the square of the difference magnifies the effect of the outliers.MSE has a wide range of applications in training models, optimizing parameters, and evaluating model performance.

4.2 Application of the algorithm

The parameters p and q in the equation represent the intensity of the convective term, i.e. the speed of motion in the fluid or the intensity of the fluid flow. In the equation, p*(du/dx) and q*(du/dy) represent the effect of the velocity gradient on the fluid in the x and y directions. When the values of p and q are larger, it means that the fluid is flowing faster and the convective term is more significant in the equation. Taking these two parameters to a value of 2, for example, can make the effect of the convective term in the equation more significant to simulate the fast flow or high-speed transport of the fluid. This allows the equations to be more consistent with the actual behavior of the fluid in a percolation-driven application. In addition, the parameter a affects the strength of the diffusion term in the equation, i.e., the second-order derivative term in the convective diffusion equation. As can be seen by the analytical solution $u(x, y) = e^{-a((x-0.5)^2+(y-0.5)^2)} + 0.5$, the parameter a determines the decay rate of the exponential function in the analytical solution. A larger value of a makes the exponential function decay faster in space, indicating that the strength of the diffusion term is higher, i.e., the change of gradient is smoothed out faster in space. a value of 5 can make the influence of the diffusion term in the equation stronger to simulate the rapid diffusion and smoothing of the fluid in space. By choosing appropriate parameter values, the strength and influence of the terms in the equation can be adjusted to better reflect the physical phenomena and behavior in real applications. In seepage drive applications, the parameter values are usually adjusted and optimized based on specific problems and experimental data to obtain more accurate simulation results and predictive power.

Combining the significance of the parameters of the above equation, the PINN method was tested to solve this convective diffusion equation by first setting the learning rate to 0.001 using the Adam optimizer and iterating 1000 times until convergence.

2023@ The American Computational Science Society. All rights reserved.

Figure 2 shows the loss function versus the mean absolute error, and the image "Loss" represents the variation of the loss function (mean square error) during the training process. The horizontal axis is the number of iterations (Epochs) and the vertical axis is the value of the mean squared error. By observing the change of the loss function, we can understand the convergence of the model during the training process; the image "Error" indicates the change of Mean Absolute Error during the training process. The horizontal axis is the number of iterations (Epochs), and the vertical axis is the value of Mean Absolute Error. By observing the change of Mean Absolute Error, the prediction accuracy of the model can be evaluated, see Figure 2.

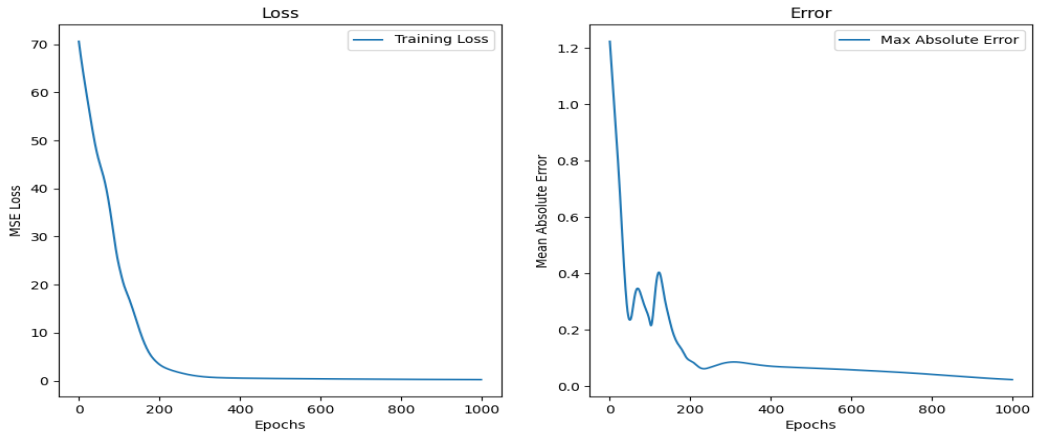

Figure 2. Loss versus error graph

As can be seen from the above table, with the increase in the number of data points, it can be seen that as the number of iterations increases exponentially, the value of the loss function through the mean square loss function shows a significant decreasing trend when the number of iterations is 0-200, after which it gradually tends to stabilize. The mean absolute error also shows an overall decreasing trend with a clear structure. The combination of the two comparisons shows that the two-dimensional convective diffusion equation is solved very well with this PINN method, and when the number of data points is 1000, a comprehensive observation of the above two two-dimensional images

It can be seen that there is the best solution effect, in this case the comparison between the true and predicted values is shown in Figure 3.

The following images represent the 3D visualization results, image "True solution" is the 3D visualization result of the true solution, showing the variation of the true solution with x and y; image "Predicted solution" is the 3D visualization result of the model predicted solution, showing the variation of the model predicted solution with x and y; image "Absolute error" is the 3D visualization result of the absolute error between the true solution and the model predicted solution, showing the variation of the error with x and y. is the 3D visualization of the absolute error between the real solution and the predicted solution, showing the variation of the error with x and y. See Figure 3.

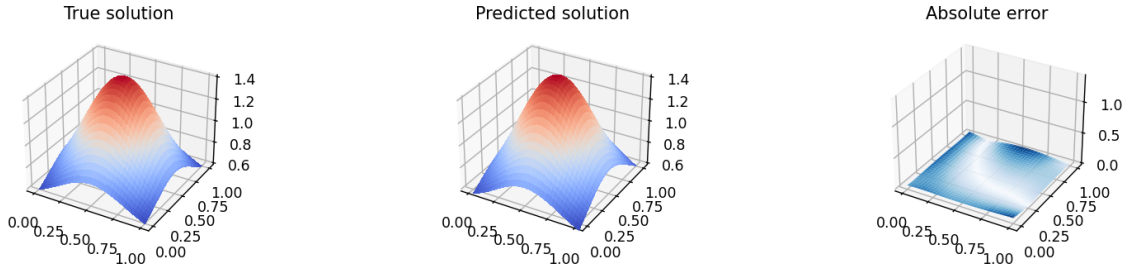

Figure 3. PINN-based learning results

2023@ The American Computational Science Society. All rights reserved.

It can be seen that the predicted value of solving this two-dimensional convective diffusion equation based on the PINN method is extremely similar to the true value of this equation, and the absolute error is close to 0. Therefore, the number of iterations is 1000 has excellent solution effect. Next, for the data points are varied to observe the changes of the maximum error value, average error value, loss function value, and running time value, and the results of the analysis are shown in Table 2 and Figure 4.

Table 2: Comparison of results

num_epochs	Max Error	Mean Error	Loss	Run time
500	0.0813	0.0334	0.02820	5.8268
1000	0.0580	0.0095	0.0154	8.7601
1500	0.0420	0.0086	0.0121	17.6619
2000	0.0405	0.0075	0.0109	32.3458

The comparison of the results in Table 2 shows that the iterative data points of the code test are spaced at 500, and the growth test process from 500 to 2000 shows that the maximum error value, average error value, and loss function value are gradually decreasing, while the running time is gradually increasing, and by observation, num_epochs=1000 can be taken as the best training result, because after this The maximum error value, average error value, and loss function values do not change significantly with the number of data points in the test, and the total time to run the code is moderate.

Figure 4 shows the true and predicted solutions of the two-dimensional convective diffusion equation for data points 500, 1000, 1500, and 2000, respectively, and shows the variation of the true and predicted solutions with x, respectively, to visualize the details of Table 2. The image "Absolute error - XZ Plane" shows the absolute error between the real solution and the predicted solution of the model for each value of data points, respectively, as shown in Fig. 4, which shows the variation of the error with x.

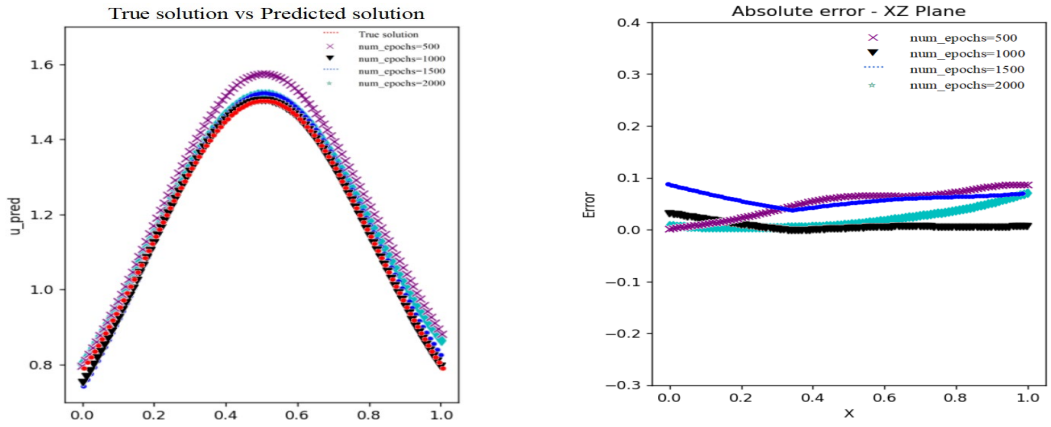

Figure 4. Contrast analysis chart

2023@ The American Computational Science Society. All rights reserved.

When the number of data points is 500, the predicted solution differs greatly compared with the real solution, and when the number of data points is taken as 1500 and 2000, the highest points of the images do not coincide with the real values, while when the number of data points is taken as 1000, it can be observed that the two-dimensional images of the real and predicted solutions have the highest coincidence, so it can be obtained that the number of data points is 1000 has the best solution effect. Meanwhile, the absolute error is stable between 0 and 0.1, which indicates that the results of the PINN method are in good agreement with the real solution of this two-dimensional convective diffusion equation.

5. Conclusion

This paper is based on the PINN solution method for solving the two-dimensional convective diffusion equation, which can be used to solve the material transport problem in seepage driven applications by choosing appropriate parameter values. The PINN method is grid-independent and has good adaptability, flexibility and practicality, and it does not require a discretized grid, which reduces the complexity of model construction and solution. By solving this equation and obtaining the analytical solution, detailed information about the material transport behavior in the seepage drive process can be obtained, which provides an important reference for the design and optimization of seepage drive applications and helps to accurately describe the fluid motion and material transport processes in porous media. Further research is needed for solving other types of convective diffusion equations.

Acknowledgments

The financial support was provided by the National Natural Science Foundation of China(NNSFC) through grant Nos 12172186,, 11772166.

References

[1] Nipho Wansophark, Pramote Dechaumphai. Streamline upwind finite element method using 6-node triangular element with adaptive remeshing technique for convective-diffusion problems[J].Applied Mathematics and Mechanics(English Edition),2008(11):1439-1450.

[2] Niphon Wansophark, Pramote Dechaumphai. Streamline upwind finite element method using 6-node triangular element with adaptive remeshing technique for convective-diffusion problems[J]. Applied Mathematics and Mechanics(English Edition),2008(11):1439-1450.

[3] Ali Umair, Iqbal Azhar, Sohail Muhammad, Abdullah Farah Aini, Khan Zohaib. compact implicit difference approximation for time-fractional diffusion- wave equation[J]. Alexandria Engineering Journal,2022,61(5).

[4]] Patil P.M., Benawadi Sunil, Sheremet Mikhail A. The unsteady nonlinear convective flow of a nanofluid due to impulsive motion: Triple diffusion and magnetic effects[J]. Thermal Science and Engineering Progress,2022,35.

[5] WANG Boqian, WANG Dingxi, RAHMATI Mohammad, HUANG Xiuquan. Revisiting the space-time gradient method: A time-clocking perspective, high order difference time discretization and comparison with the harmonic balance method[J]. Chinese Journal of Aeronautics,2022,35(11).

[6] Cen Dakang, Wang Zhibo. Time two-grid technique combined with temporal second order difference method for two-dimensional semilinear fractional sub diffusion equations[J]. Applied Mathematics Letters,2022(prepublish).

[7] Labidi Samira, Omrani Khaled. A new conservative fourth-order accurate difference scheme for the nonlinear Schrödinger equation with wave operator [J]. Applied Numerical Mathematics,2022,173.

[8] Daosheng XU, Dehui CHEN,Kaixin WU. Properties of High-Order Finite Difference Schemes and Idealized Numerical Testing[J].Advances in Atmospheric Sciences,2021,38(04):615-626.

[9] Zijian Zhou, Zhenya Yan. Deep learning neural networks for the third-order nonlinear Schr? dinger equation: bright solitons, breathers, and rogue waves[J]. Communications in Theoretical Physics,2021,73(10):57-65.

[10] Jiaheng Li, Biao Li. Solving forward and inverse problems of the nonlinear Schr?dinger equation with the generalized PT-symmetric Scarf-II potential via PINN deep learning[J].Communications in Theoretical Physics,2021,73(12):5-17.

[11] Hui Pang, Longxing Wu, Jiahao Liu, Xiaofei Liu, Kai Liu. Physics-informed neural network approach for heat generation rate estimation of lithium-ion battery under various driving conditions[J]. Journal of Energy Chemistry,2023,78(03):1-12.

[12] Chengping Rao, Hao Sun, Yang Liu. Physics-informed deep learning for incompressible laminar flows[J]. Theoretical & Applied Mechanics Letters. 2020,10(03):207-212.

[13] Data-driven parity-time-symmetric vector rogue wave solutions of multi-component nonlinear Schr?dinger equation[J].Chinese Physics B,2022,31(06):169-176.

[14] M.HAMID, M.USMAN, Zhenfu TIAN. Computational analysis for fractional characterization of coupled convection-diffusion equations arising in MHD fows[J].Applied Mathematics and Mechanics(English Edition),2023,44(04):669-692.

[15] ZHANG Shaotong, JIA Yonggang, WEN Mingzheng, WANG Zhenhao, Zhang Yaqi, ZHU Chaoqi, Li Bowen, LIU Xiaolei. Vertical Migration of Fine-Grained Sediments from Interior to Surface of Seabed Driven by Seepage Flows-'Sub-Bottom Sediment Pump Action'[J]. Journal of Ocean University of China,2017,16(01):15-24.

[16] WANG Yang, YU Haiyang, ZHANG Jia, FENG Naichao, CHENG Shiqing. Transient pressure analysis of

2023@ The American Computational Science Society. All rights reserved.

polymer flooding fractured wells with oil-water two- phase flow[J]. Petroleum Exploration and Development,2023,50(01):175-182.

[17] Tadmor E. The numerical viscosity of entropy stable schemes for systems of conservation laws I[J]. Mathematics of Computation,1987,49(179): 91-103.

[18] Roe P L. Affordable, entropy-consistent Euler flux functions II: entropy production at shocks[J].Journal of Computational Physics,2009,228(15). 5410-5436.

[19] Sahoo Mrutyunjaya, Chakraverty Snehashish. Sawi Transform Based Homotopy Perturbation Method for Solving Shallow Water Wave Equations in Fuzzy Environment[J]. Mathematics,2022,10(16).

[20] Ismail F, Roe P L. Affordable, entropy-consistent Euler flux functions II: Entropy production at shocks[J]. Journal of Computational Physics,2009. 228(1): 5410-5436.

[21] Fjordholm U S, Mishra S, Tadmor E. Arbitrarily high-order accurate entropy stable essentially non-oscillatory schemes for systems of conservation laws[J]. Siam Journal on Numerical Analysis,2012,50(2): 544-573.

[22] Levy D, Puppo G, Russo G. A fourth-order central WENO scheme for multidimensional hyperbolic systems of conservation laws[J].SIAM Journal on Scientific Computing, 2002,24(2): 480--506.

[23] Han J, Arnulf J, Weinan E. Solving high dimensional partial differential equations using deep learning[J]. Proceedings of the National Academy of Sciences,2018,115(34): 8505- 8510.

[24] Sirignano J, Spiliopoulos K. DGM: A deep learning algorithm for solving partial differential equations[J]. Journal of Computational Physics,2018,375: 1339-1364.

[25] Sun L N, Gao H,Pan S W, etal. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data[J].Computer Methods in Applied Mechanics and Engineering,2020,361(C): 112732

2023@ The American Computational Science Society. All rights reserved.