Ukrainian Journal of Physical Optics
SPECIAL ISSUE 5, 2024

PREDICTION OF NONDEGENERATE SOLITONS AND PARAMETERS IN NONLINEAR BIREFRINGENT OPTICAL FIBERS USING PHPINN AND DEEPONET ALGORITHMS

SU-YONG XU[1,2], AO-CHENG YANG[1], QIN ZHOU[1,2,3]

[1] Research Group of Nonlinear Optical Science and Technology, Research Center of Nonlinear Science, School of Mathematical and Physical Sciences, Wuhan Textile University, Wuhan 430200, China, qinzhou@whu.edu.cn
[2] College of Optical, Mechanical and Electrical Engineering, Zhejiang A&F University, Lin'an 311300, China
[3] State Key Laboratory of New Textile Materials and Advanced Processing Technologies, Wuhan Textile University, Wuhan 430200, China

Received: 05.04.2024

Abstract. This paper uses parallel hard-constraint physics-informed neural networks to investigate the prediction of nondegenerate soliton and estimate parameters for the coupled nonlinear Schrödinger's equation. Based on our previous analytical results, three types of nondegenerate solitons have been predicted in the forward problem under the corresponding initial and boundary conditions. In the inverse problem, when employing pure data as the training set, the relative errors in predicting the system's parameters of group velocity dispersion β_2 and Kerr nonlinearity γ are both less than 1%. Moreover, upon introducing a 5% noise level to the training set, the relative errors for β2 and γ remain below 3%. Additionally, we introduce for the first time the application of Deep Operator Networks (DeepONet) to predict nondegenerate soliton, reducing relative L2 error to 10^{-3} and achieving a speedup of approximately 10^3 times higher compared to the phPINN method. This demonstrates the efficacy of operator learning methods in addressing nonlinear optical problems.

Keywords: nondegenerate solitons, coupled nonlinear Schrödinger's equation, phPINN, DeepONet
UDC: 535.32
DOI: 10.3116/16091833/Ukr.J.Phys.Opt.2024.S1137

1. Introduction

Optical solitons, also known as temporal solitons, result from a balance between dispersion and nonlinearity in optical fibers [1-3]. Specifically, bright solitons are formed in the anomalous dispersion region of the optical fiber, while dark solitons are formed in the normal dispersion region. When an intense picosecond pulse is transmitted in a single-mode optical fiber, its dynamics are described by the standard or higher-order nonlinear Schrödinger's equation. However, when optical waves with different modes, such as the same polarization state at different wavelengths or different polarization directions at the same wavelength, are transmitted together in the optical fiber, we need to extend this model to the coupled nonlinear Schrödinger's equation (CNLSE) [4-6]. Recently, through the integrable theoretical methods, including the Darboux transformation and Hirota's bilinear methods, many new structures of fiber solitons have been reported, such as the nondegenerate solitons in birefringent fibers, double-valley dark solitons in highly nonlinear optical fiber, solitons on the nonvanishing background and so on [7-10].

Artificial intelligence has emerged as a prominent topic in recent years, with deep learning finding widespread applications in engineering and scientific computing. Notably, the concept of scientific machine learning has garnered considerable attention [11-19]. Among the foremost methodologies in this domain are physics-informed neural networks (PINN) [20] and neural operators. Physics-informed neural network (PINN) is adept at tackling both forward and inverse problems associated with nonlinear partial differential equations (PDEs). Its fundamental premise involves integrating physical constraints into the neural network training process. PINN can approximate unknown solution functions by incorporating known physical equations as constraints and glean insights into system behavior from available data.

On the other hand, deep neural operators can learn nonlinear mappings across infinite-dimensional function spaces using deep neural networks [21]. Serving as promising surrogate solvers for real-time prediction of partial differential equations, deep neural operators like deep operator networks (DeepONet) offer a novel simulation paradigm in science and engineering. Researchers have utilized PINN to address challenges in nonlinear optics through adjustments in network architecture, expressions of loss functions, and methods for sampling configuration points [22-27]. Our previous work proposed the parallel hard-constraint physics-informed neural networks (phPINN) and successfully improved the accuracy and efficiency of solving the NLS-Maxwell-Bloch equation [28]. DeepONet [29], the first neural operator, has performed well in building surrogate models for many PDEs. For example, DeepONet has been applied in multiscale bubble dynamics [30], boundary layer instabilities [31], solar–thermal systems forecasting [32], and fast multiscale modeling [33-37]. However, research on DeepONet in solitons has not been reported yet. This study will use deep learning methods, including parallel hard-constraint physics-informed neural networks (phPINN) and DeepONet, to investigate nondegenerate solitons.

In our previous work [9], the analytical optical nondegenerate solitons in a birefringent fiber with a 35-degree elliptical angle are reported. The following CNLSE is considered:

$$\begin{aligned} u_z + \beta_1 u_t + \frac{\mathrm{i}}{2}\beta_2 u_{tt} - \mathrm{i}\beta_0 u - \mathrm{i}\gamma(|u|^2 + |v|^2)u = 0, \\ v_z + \beta_1 v_t + \frac{\mathrm{i}}{2}\beta_2 v_{tt} - \mathrm{i}\beta_0 v - \mathrm{i}\gamma(|u|^2 + |v|^2)v = 0, \end{aligned} \tag{1}$$

where $u(z,t)$ and $v(z,t)$ are the two polarization components of light wave, β_0 represents the wave number difference between two modes, β_1 is the reciprocal of group velocity, β_2 originates from group velocity dispersion, and γ is the Kerr nonlinearity.

2. The phPINN and DeepONet methods

In this work, we investigated the accuracy of phPINN and DeepONet in predicting nonlinear optical problems. A well-trained phPINN network represents the equation solution under a certain initial and boundary condition, while a well-trained DeepONet can quickly obtain the equation solution corresponding to any initial condition covered by the training set. Sections 2.1 to 2.3 provide detailed information on data, neural network architecture, loss functions, and training environment.

2.1. The phPINN for soliton evolution and equation parameters prediction

phPINN aims to improve the performance of neural networks in predicting soliton solutions and parameter estimation in nonlinear fiber optic systems by improving the design of loss

functions and network architecture. The phPINN network effectively predicted the solutions of the NLS-MB equation, including single solitons, multiple solitons, and odd waves, and successfully solved the parameter estimation problem related to the NLS-MB equation. We employ four fully-connected neural networks and hard-constraint initial conditions for all examples.

We adopt the CNLSE as the physical constraint to construct the phPINN above. The u and v in Eq. (1) are the complex solutions of z and t, which will be determined subsequently, necessitating the separation of the real and imaginary components. We obtain the model corresponding to Eq. (1) for the phPINN as:

$$\begin{cases} f_1 \coloneqq \gamma \mathrm{Im}\{u\}(\mathrm{Im}\{u\}^2 + \mathrm{Re}\{u\}^2) + \gamma \mathrm{Im}\{u\}(\mathrm{Im}\{v\}^2 + \mathrm{Re}\{v\}^2) \\ \quad + \mathrm{Im}\{u\}\beta_0 + \mathrm{Re}\{u_t\}\beta_1 - \dfrac{\mathrm{Im}\{u_{tt}\}\beta_2}{2} + \mathrm{Re}\{u_z\}, \\ f_2 \coloneqq -\gamma \mathrm{Re}\{u\}(\mathrm{Im}\{u\}^2 + \mathrm{Re}\{u\}^2) - \gamma \mathrm{Re}\{u\}(\mathrm{Im}\{v\}^2 + \mathrm{Re}\{v\}^2) + \\ \quad \mathrm{Im}\{u_t\}\beta_1 + \dfrac{\mathrm{Re}\{u_{tt}\}\beta_2}{2} - \mathrm{Re}\{u\}\beta_0 + \mathrm{Im}\{u_z\}, \\ f_3 \coloneqq \gamma \mathrm{Im}\{v\}(\mathrm{Im}\{u\}^2 + \mathrm{Re}\{u\}^2) + \gamma \mathrm{Im}\{v\}(\mathrm{Im}\{v\}^2 + \mathrm{Re}\{v\}^2) \\ \quad + \mathrm{Im}\{v\}\beta_0 + \mathrm{Re}\{v_t\}\beta_1 - \dfrac{\mathrm{Im}\{v_{tt}\}\beta_2}{2} + \mathrm{Re}\{v_z\}, \\ f_4 \coloneqq -\gamma \mathrm{Re}\{v\}(\mathrm{Im}\{u\}^2 + \mathrm{Re}\{u\}^2) - \gamma \mathrm{Re}\{v\}(\mathrm{Im}\{v\}^2 + \mathrm{Re}\{v\}^2) - \mathrm{Re}\{v\}\beta_0 \\ \quad + \mathrm{Im}\{v_t\}\beta_1 + \dfrac{\mathrm{Re}\{v_{tt}\}\beta_2}{2} + \mathrm{Im}\{v_z\}, \end{cases} \tag{2}$$

where Re{·} and Im{·} represent the real and imaginary parts of the respective quantities.

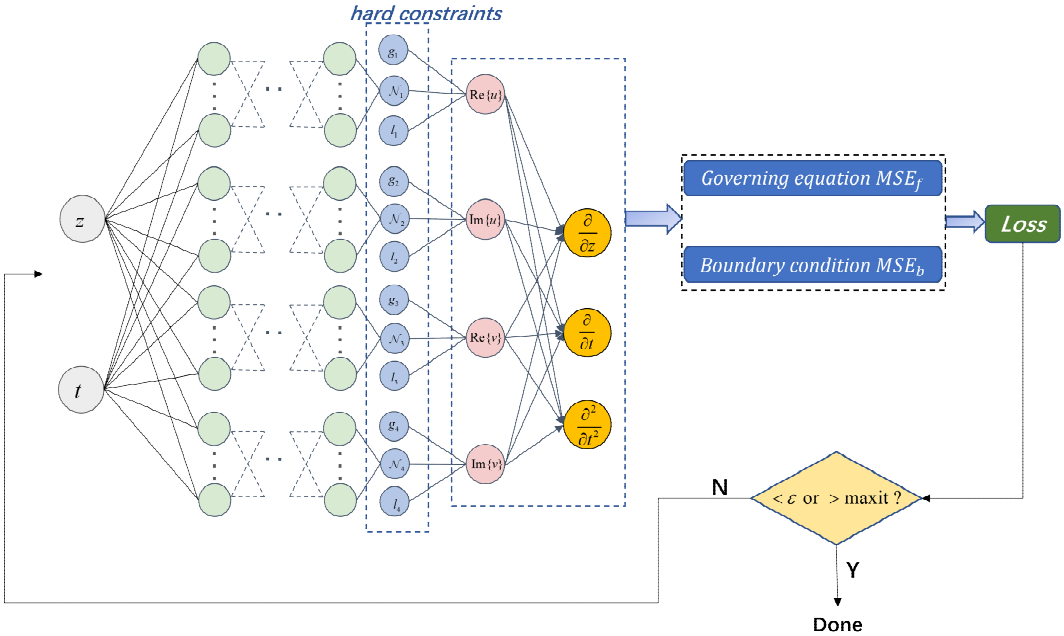

Fig. 1. The framework structure of the phPINN for the CNLSE system.

Fig. 1 shows the training process of the phPINN architecture. We first initialize the network parameters to obtain the initial estimates of the corresponding variables Re{u}, Im{u}, Re{v}, and Im{v} in Eq. (1). Give the initial estimation, we subsequently formulate a

loss function to minimize the PDE residuals, as well as the errors between the ground truth and the approximation in the boundary. Then iteratively update the network parameters until the total loss converges.

2.2. DeepONet

The methodologies for solving differential equations can be broadly categorized into several classes. One category encompasses traditional solvers, such as those based on finite difference, finite element, and finite volume methods, including Jacobi iteration, multigrid iterations, and spectrum discontinuous Galerkin methods. These methods conventionally discretize the problem onto a grid or a finite basis and subsequently solve a linear system or iterate over it to obtain the desired solution. In contrast, there exists a paradigm known as operator learning, which represents a surrogate model aimed at approximating the input-output relationship or operator function in a simplified or lower-dimensional manner.

Operator learning entails learning the input-output relationship defined by a given differential equation. It involves capturing relationships between functions over time steps or coefficients and their accumulations called operators. Like traditional supervised learning, operator learning typically involves collecting datasets comprising various medium conditions or initial conditions and corresponding solutions obtained through solvers or experimental observations. Subsequently, supervised learning techniques are applied to learn these relationships. Contrary to traditional methods, which discretize inputs and outputs into grids and solve them at each point using linear methods, operator learning resembles a statistical inference process that finds correspondences between input function spaces and output function spaces. This facilitates rapid prediction of results for new inputs, offering significant speedup compared to traditional approaches, often by orders of magnitude.

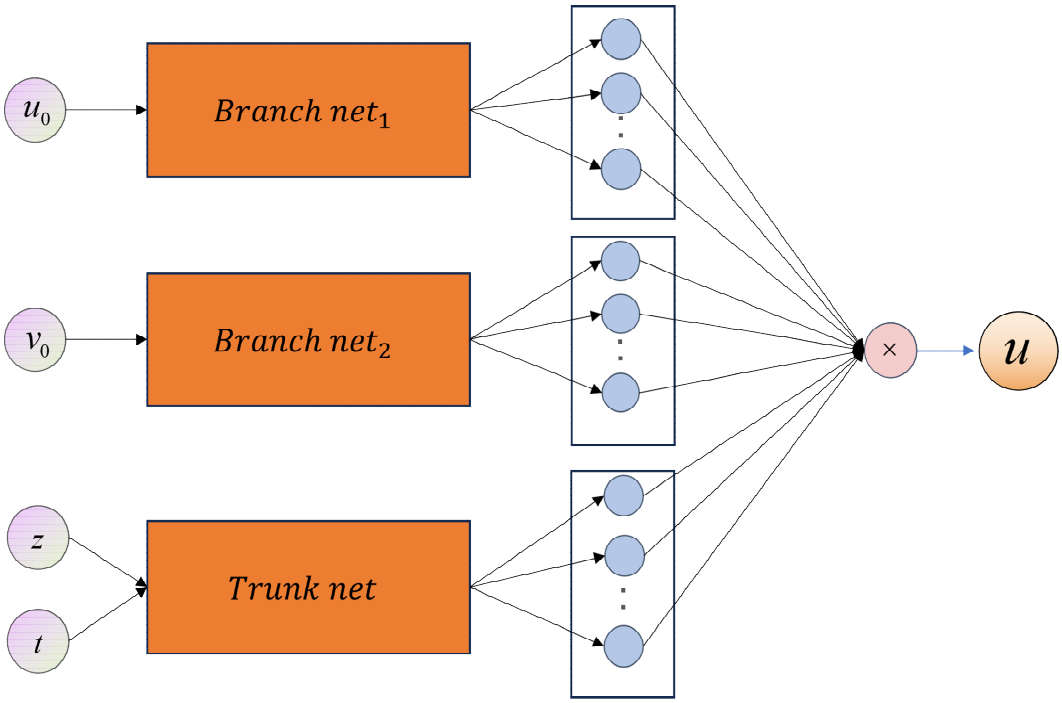

Fig. 2. The framework structure of DeepONet for the CNLS system.

The concept of neural operators draws inspiration from methods like Green's functions, commonly used in linear PDEs, where solutions are expressed in terms of integral functions. However, for more challenging nonlinear problems, linear transformations alone are insufficient to capture the behavior of the underlying operators. Therefore, linear and nonlinear operations are combined to approximate any continuous but nonlinear operators effectively. Early works in this domain employed graph neural networks to transform input functions into graph structures and then used graph neural network techniques to perform integrations, effectively treating them as message-passing operations. The DeepONet has been introduced by Lu et al. in 2019. The operators were decomposed into branch and trunk networks, enabling the representation of operators as tensor products of these networks' outputs. Fig. 2 shows the network structure of the DeepONet used in this article. *Branch net*$_1$ and *Branh net*$_2$ input the initial conditions of *u* and *v*, while the *Trunk net* inputs coordinate. All the branch nets and the trunk net have the same number of outputs, merged via the Hadamard product and a subsequent summation. The entire network outputs the solution of the equation, i.e., *u* or *v*.

2.3. Training data and network environment

In supervised learning, training data is crucial for training neural networks. For this problem, we can obtain training data from the exact solution. In addition, we chose Xavier initialization and hyperbolic tangent (tanh) as activation functions. This article uses a pseudorandom generator with a PCG-64 algorithm to generate residual training points. All the code in this article was built using the DeepXDE library [38], based on Python 3.10 and Tensorflow 2.10. All numerical experiments reported in this article were run on a computer with an Intel (R) Xeon (R) W-2265 CPU @ 3.50GHz processor, 64GB memory, and 10GB Nvidia GeForce RTX 3080 graphics card.

3. The forward and inverse problems by phPINN algorithm

In this chapter, we will delve into the application of the phPINN algorithm in CNLSE. We aim to solve both forward and inverse problems, namely, accurate prediction of soliton solutions and accurate estimation of system parameters.

3.1. Prediction of soliton molecules

Considering Eq. (1) with the parameter values γ=-4, β_0=β_1=2 and β_2=1, the explicit soliton molecules solution can be found in Eq. (4) of reference [9]. Specifically, we select parameters $k_1=\frac{1}{2}+\frac{1}{4}i$, $l_1=0.5001+\frac{1}{4}i$, $\alpha_1^{(1)}=\frac{1}{6}-i$, and $\alpha_1^{(2)}=\frac{1}{2}+i$. To obtain the original training dataset for the specified initial and boundary conditions, the spatial region [-20, 20] and the temporal region [-20, 30] are discretized into 512 points, and the exact solution of the soliton molecules is discretized accordingly. These data will be used to calculate the relative L2 error. Furthermore, a smaller training set containing boundary data is generated by randomly selecting the boundary data points N_b=1000 from the original dataset. The initial value of the number of PDE residual training points is N_f=25000. With 30000 iterations using the Adam optimizer and an additional 9553 iterations using the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) optimizer, the phPINN framework successfully learns the soliton molecules utilizing four subnetworks. The network achieves the relative L2 error of *u* as 9.026×10^{-4} and *v* as 9.193×10^{-4}, with a total of 39563 iterations and a training time of 2704 seconds.

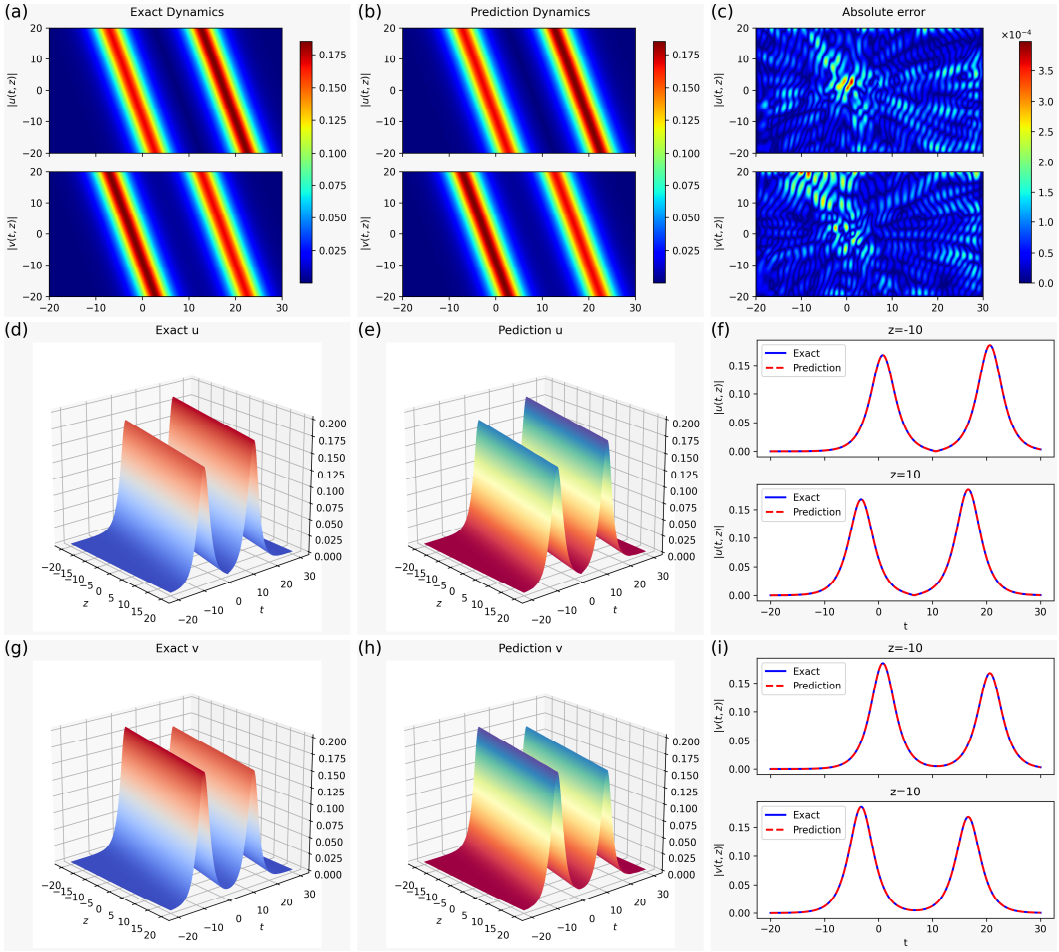

Fig. 3. Prediction results of soliton molecules. (a) and (b) show the exact and predicted dynamic evolution of soliton molecules, respectively. (c) present the point-by-point absolute error between the predicted and exact solutions. (d), (e), (g), and (h) show the 3D stereograms of the true and predicted solutions of soliton molecules, respectively. (f) and (i) show the comparison between the exact (solid curve) and predicted (dash curve) solutions of u and v at z=-10 and z=10, respectively.

Fig. 3 illustrates the deep learning results of soliton molecules based on the CNLSE using the phPINN. The phPINN architecture effectively resolves the local features of soliton molecules' dynamics. The predicted solution closely matches the exact solution. The absolute errors were up to 10^{-4} orders for the u and v fields.

3.2. Prediction of the symmetric double-double-hump solitons

Selecting parameters $k_1=\frac{1}{3}+\frac{1}{4}i$, $l_1=\frac{1}{2}+\frac{1}{4}i$, $\alpha_1^{(1)}=\frac{2}{3}-\frac{1}{3}i$, and $\alpha_1^{(2)}=\frac{1}{2}+i$, the explicit symmetric double-double-hump solitons solution can be found in reference [9]. Using the same scheme as soliton molecules, except for the temporal region being changed to [-30,30] to construct the network and training set, the phPINN framework successfully learns the symmetric double-double-hump solitons utilizing four subnetworks. The network achieves a relative L2 error of u as 1.162×10^{-3}, and v as 9.649×10^{-4}, with a total of 38478 iterations and a training time of 1743 seconds.

Fig. 4 presents the outcomes of deep learning applied to symmetric, symmetric double-double-hump solitons. The phPINN architecture effectively resolves the local features of the dynamics of the symmetric double-double-hump solitons. The order of absolute errors was up to 10^{-4} for the u and v fields. A close examination reveals the point-by-point absolute error between the predicted and exact solutions, with a peak value of merely 4×10^{-4}.

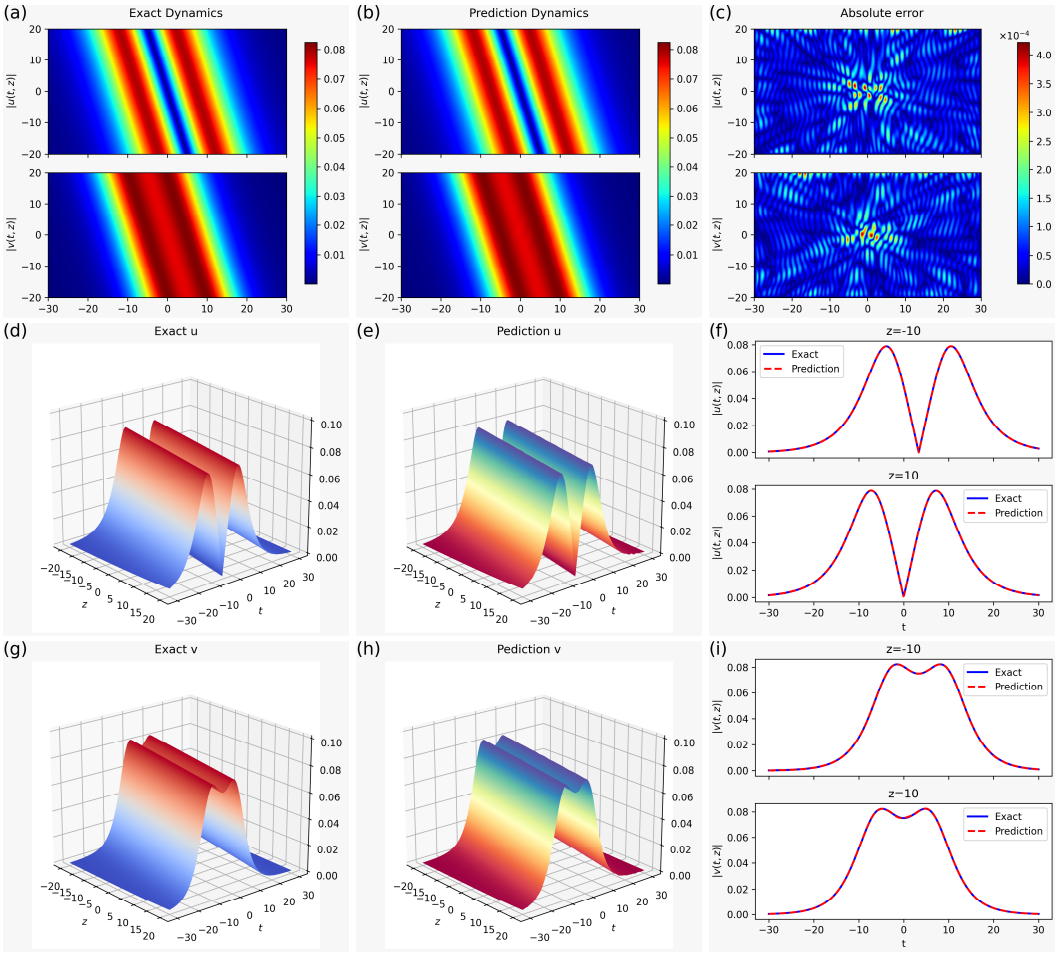

Fig. 4. Prediction results of the symmetric double-double-hump solitons. (a) and (b) show the exact and predicted dynamic evolution of the symmetric double-double-hump solitons, respectively. (c) present the pointwise error between the predicted and exact solutions. (d), (e), (g), and (h) show the 3D stereograms of the true and predicted solutions of soliton molecules, respectively. (f) and (i) show the comparison between the exact (solid curve) and predicted (dash curve) solutions of u and v at z=-10 and z=10, respectively.

3.3. Prediction of the symmetric double-hump-single-hump solitons

In this section, we consider predicting symmetric double-hump-single-hump solitons. Using the same scheme as symmetric double-double-hump solitons to construct the network and training set, the phPINN framework successfully learned symmetric double-hump-single-hump solitons through 30000 iterations using the Adam optimizer and 8478 iterations using the L-BFGS optimizer. The network achieved a relative L2 error u of 1.162×10^{-3} and v of 9.649×10^{-4}, with 38478 iterations and a training time of 1743 seconds. Fig. 5 shows the phPINN architecture effectively resolves the local features of symmetric double-hump-

single-hump solitons dynamics. The absolute errors were up to 10^{-4} order for the u and v fields, with the maximum pointwise absolute error reaching only 4×10^{-4}.

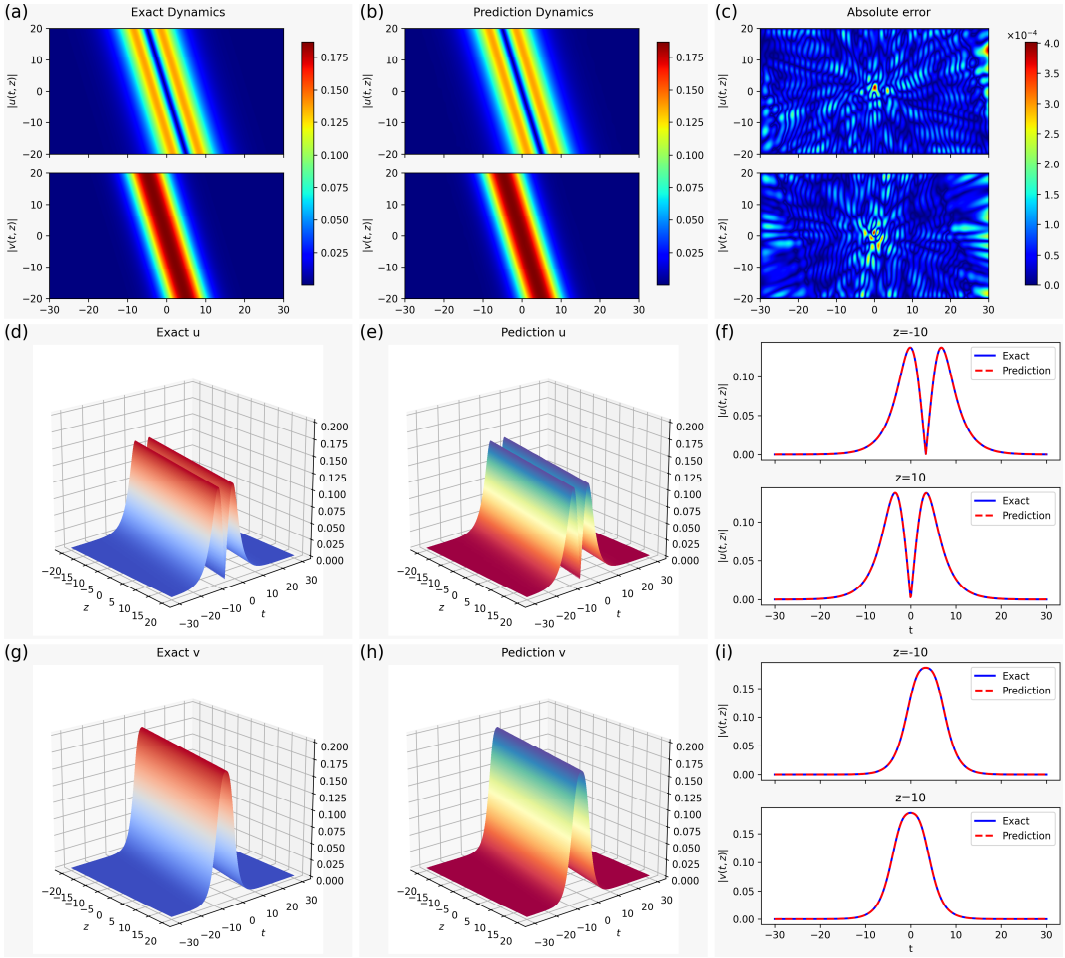

Fig. 5. Prediction results of the symmetric double-hump-single-hump solitons. (a) and (b) show the exact and predicted dynamic evolution of the symmetric double-hump-single-hump solitons, respectively. (c) present the pointwise error between the predicted and exact solutions. (d), (e), (g), and (h) show the 3D stereograms of the true and predicted solutions of soliton molecules, respectively. (f) and (i) show the comparison between the exact (solid curve) and predicted (dash curve) solutions of u and v at z=-10 and z=10, respectively.

3.4. Prediction of equation parameters

In this section, our attention is directed towards addressing the inverse problem associated with CNLSE systems, wherein the task entails discerning the parameters of the CNLSE system model by utilizing known solutions as the training dataset. To ascertain the unknown parameters β_2 and γ in Eq. (1), we employ the phPINN framework previously utilized. 20000 data points are randomly sampled from the original symmetric double-hump-single-hump solitons solution to form the training dataset. The residual equation is evaluated utilizing N_f=10000 points generated through the pseudorandom method. The unknown parameters β_2 and γ are initialized with arbitrary values upon preparing the training dataset. After 50000 iterations using the Adam optimizer and additional L-BFGS iterations, all the learnable parameters of the phPINN model are optimized, and the loss function is adjusted to

facilitate the prediction of the unknown parameters β_2 and γ. The relative error of the unknown parameters is defined as:

$$\mathrm{RE} = \left| \frac{\hat{\delta} - \delta}{\delta} \right| \times 100\%, \tag{3}$$

where $\hat{\delta}$ and δ represent the predicted and true values, respectively.

Fig. 6 shows our prediction results. Fig. 6a describes the dependence of unknown parameters β_2 and γ on the iterations number, with shadows representing an error range of ± 10%, dashed lines representing true values, and solid lines representing learned parameters. The final relative error of β_2 is 0.031%, and the final relative error of γ is 0.502%. Afterward, we added 5% noise to the solution and continued the experiment, as shown in Fig. 6b. The final relative error of β_2 is 0.124%, and the final relative error of γ is 2.188%.

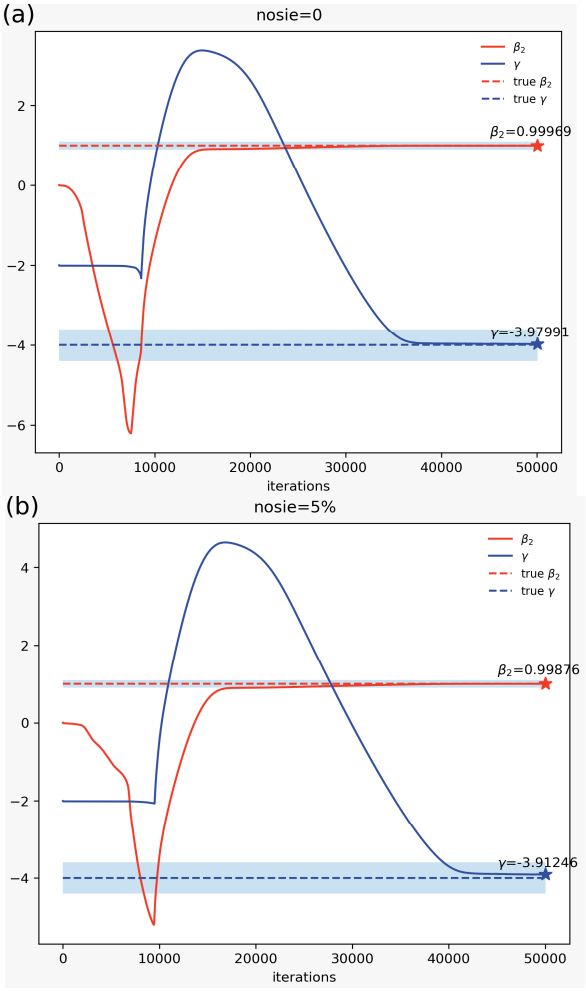

Fig. 6. phPINN parameter estimation results. The dependence of unknown parameters β_2 and γ on the iterations number under noise levels of (a) 0% and (b) 5%, respectively.

3.5. Predicting nondegenerate solitons by DeepONet method

To test the DeepONet model's ability to predict nondegenerate optical solitons in nonlinear birefringent fibers, we tested its prediction on a series of different physical parameters. Liu's work used bilinear methods to derive the following expressions for u and v [9]:

$$u=\frac{\alpha_0\left(1+\frac{2\mathrm{e}^{\frac{\left(l_1+\overline{l_1}\right)\left(2t-i\left(l_1-\overline{l_1}\right)z\beta_2\right)}{2}}\left(-k_1+l_1\right)\alpha_1\overline{\alpha_1}\gamma}{5\left(k_1+\overline{l_1}\right)\left(l_1+\overline{l_1}\right)^2\beta_2}\right)\mathrm{e}^{-\frac{i\left(-2t\left(-ik_1\beta_2+\beta_1\right)+z\left(\beta_1^2+\beta_2\left(\beta_2k_1^2-2\beta_0\right)\right)\right)}{2\beta_2}}}{\left[\begin{array}{l}1+\frac{\gamma\left(-\frac{3\mathrm{e}^{\frac{\left(k_1+\overline{k_1}\right)\left(2t-i\left(k_1-\overline{k_1}\right)z\beta_2\right)}{2}}\alpha_0\overline{\alpha_0}}{5\left(k_1+\overline{k_1}\right)^2}-\frac{2\mathrm{e}^{\frac{\left(l_1+\overline{l_1}\right)\left(2t-i\left(l_1-\overline{l_1}\right)z\beta_2\right)}{2}}\alpha_1\overline{\alpha_1}}{5\left(l_1+\overline{l_1}\right)^2}\right)}{\beta_2}\\+\frac{6\mathrm{e}^{k_1t+\overline{k_1}t-\frac{ik_1^2z\beta_2}{2}+\frac{i\overline{k_1}^2z\beta_2}{2}+\frac{\left(l_1+\overline{l_1}\right)\left(2t-i\left(l_1-\overline{l_1}\right)z\beta_2\right)}{2}}\left(k_1-l_1\right)\left(\overline{k_1}-\overline{l_1}\right)\alpha_0\overline{\alpha_0}\alpha_1\overline{\alpha_1}\gamma^2}{25\left(k_1+\overline{k_1}\right)^2\left(\overline{k_1}+l_1\right)\left(k_1+\overline{l_1}\right)\left(l_1+\overline{l_1}\right)^2\beta_2^2}\end{array}\right]},$$

$$v=\frac{\alpha_1\left(1+\frac{3\mathrm{e}^{\frac{\left(k_1+\overline{k_1}\right)\left(2t-i\left(k_1-\overline{k_1}\right)z\beta_2\right)}{2}}\left(k_1-l_1\right)\alpha_0\overline{\alpha_0}\gamma}{5\left(k_1+\overline{k_1}\right)^2\left(\overline{k_1}+l_1\right)\beta_2}\right)\mathrm{e}^{-\frac{i\left(-2t\left(-il_1\beta_2+\beta_1\right)+z\left(\beta_1^2+\beta_2\left(\beta_2l_1^2-2\beta_0\right)\right)\right)}{2\beta_2}}}{\left[\begin{array}{l}1+\frac{\gamma\left(-\frac{3\mathrm{e}^{\frac{\left(k_1+\overline{k_1}\right)\left(2t-i\left(k_1-\overline{k_1}\right)z\beta_2\right)}{2}}\alpha_0\overline{\alpha_0}}{5\left(k_1+\overline{k_1}\right)^2}-\frac{2\mathrm{e}^{\frac{\left(l_1+\overline{l_1}\right)\left(2t-i\left(l_1-\overline{l_1}\right)z\beta_2\right)}{2}}\alpha_1\overline{\alpha_1}}{5\left(l_1+\overline{l_1}\right)^2}\right)}{\beta_2}\\+\frac{6\mathrm{e}^{k_1t+\overline{k_1}t-\frac{ik_1^2z\beta_2}{2}+\frac{i\overline{k_1}^2z\beta_2}{2}+\frac{\left(l_1+\overline{l_1}\right)\left(2t-i\left(l_1-\overline{l_1}\right)z\beta_2\right)}{2}}\left(k_1-l_1\right)\left(\overline{k_1}-\overline{l_1}\right)\alpha_0\overline{\alpha_0}\alpha_1\overline{\alpha_1}\gamma^2}{25\left(k_1+\overline{k_1}\right)^2\left(\overline{k_1}+l_1\right)\left(k_1+\overline{l_1}\right)\left(l_1+\overline{l_1}\right)^2\beta_2^2}\end{array}\right]} \quad (4)$$

set $k_1=\frac{1}{5}+\frac{1}{4}i$, $l_1=\frac{1}{4}+\frac{1}{4}i$, $\alpha_1^{\{1\}}=\frac{\sqrt{5}}{10}-\frac{1}{2}i$ and $\alpha_1^{\{2\}}=\frac{\sqrt{3}}{4}+\frac{3}{8}i$, we can get

$$u=\frac{\left[\begin{array}{l}4\mathrm{e}^{\frac{(40+9i)z\beta_2^2+\left(800iz+(160+200i)t\right)\beta_2-i(-800t+400z)}{800\beta_2}}\\\times\left(\sqrt{5}-5i\right)\left(7\mathrm{e}^{\frac{t}{2}+\frac{z\beta_2}{8}}\gamma+120\beta_2\right)\beta_2\end{array}\right]}{\left[2520\gamma\mathrm{e}^{\frac{t}{2}+\frac{z\beta_2}{8}}\beta_2+5400\mathrm{e}^{\frac{2t}{5}+\frac{z\beta_2}{10}}\gamma\beta_2-35\gamma^2\mathrm{e}^{\frac{9t}{10}+\frac{9z\beta_2}{40}}-4800\beta_2^2\right]}$$

$$v=\frac{\left[\begin{array}{l}15\mathrm{e}^{-\frac{i\left(zi\beta_2^2+4ti\beta_2-4t\beta_2-16z\beta_2-16t+8z\right)}{16\beta_2}}\\\times\left(2\sqrt{3}+3i\right)\left(\mathrm{e}^{\frac{2t}{5}+\frac{z\beta_2}{10}}\gamma-8\beta_2\right)\beta_2\end{array}\right]}{\left[504\gamma\mathrm{e}^{\frac{t}{2}+\frac{z\beta_2}{8}}\beta_2+1080\mathrm{e}^{\frac{2t}{5}+\frac{z\beta_2}{10}}\gamma\beta_2-7\gamma^2\mathrm{e}^{\frac{9t}{10}+\frac{9z\beta_2}{40}}-960\beta_2^2\right]} \quad (5)$$

We constructed a training and testing set consisting of the values of β_2 and γ. The specific steps are as follows: let β_2 = 5,5.5,6,6.5,7,7.5,8,8.5,9,9.5,10,15,20,25,30, 35,40,45,50,55} and γ =-{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}. Then, by permeating and combining each value of β_2 and γ into Eq. (4), we obtain 400 solutions, which serve as our training set. Similarly, the parameters β_2 that comprise the test set are randomly selected as 10 numbers between 5 and 55, and γ is randomly selected as 10 numbers between -11 and -0.1.

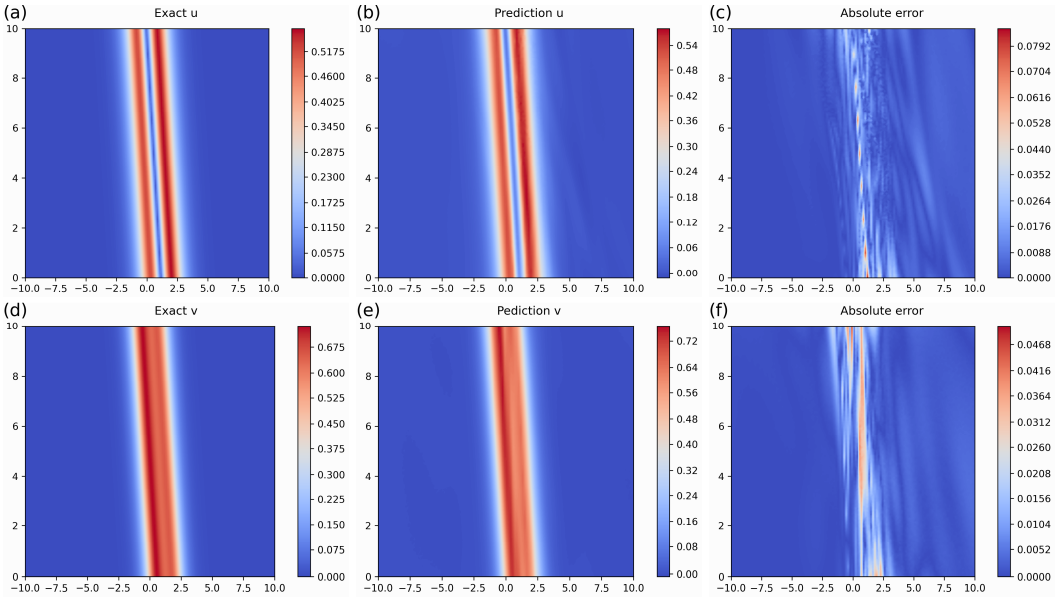

Fig. 7. The DeepONet prediction for a test sample with β_2=33.6, γ=-4.68.

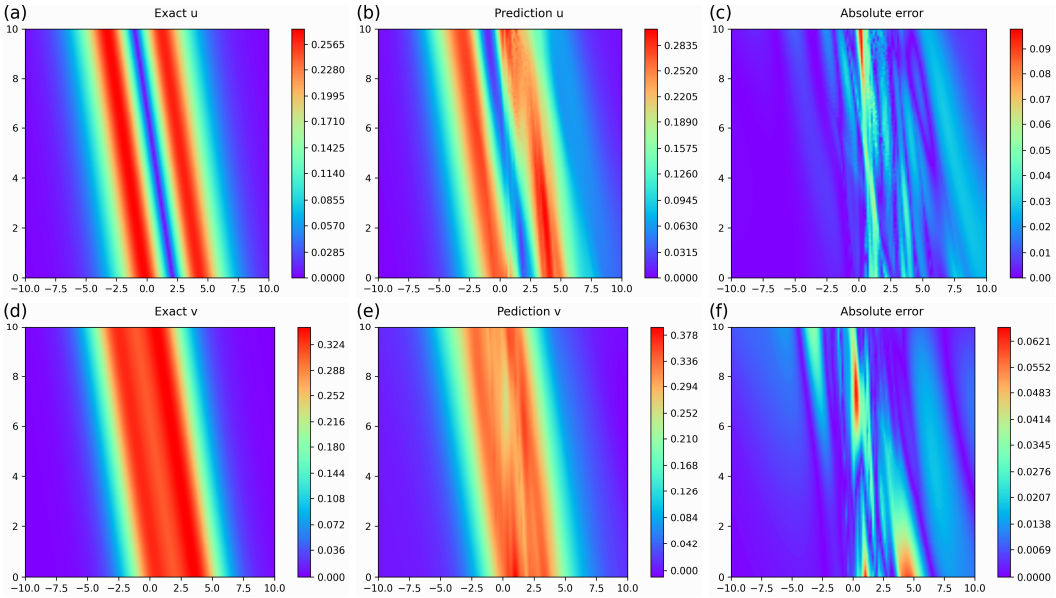

Fig. 8. The DeepONet prediction for a test sample with β_2=12.8, γ=-7.4.

As described in Section 2, the DeepONet model takes the initial conditions as input and uses the dot product coarse grid of networks *Branch net*$_1$, *Branch net*$_2$, and *Trunk net* to predict the solution of the CNLSE. Compare these predictions with actual values. The sample prediction is shown in Fig. 7 and Fig. 8. The relative L2 error of $model_u$ is 9.99×10^{-3}, and the relative L2 error of $model_v$ is 8.80×10^{-3}. Fig. 9a,b show the training process of $model_u$ and $model_v$, respectively. The relative L2 error of the two polarization components of light and the training loss convergence history are presented. We observed rapid convergences in the loss function and relative L2 error over the first 10000 iterations. The relative L2 error reached a roughly stable state after 20000 iterations. Fig. 9c shows the time and error

distribution required for the phPINN and DeepONet methods to obtain a solution. DeepONet can quickly achieve a solution with a small error, underscoring a key advantage of operator learning methodologies.

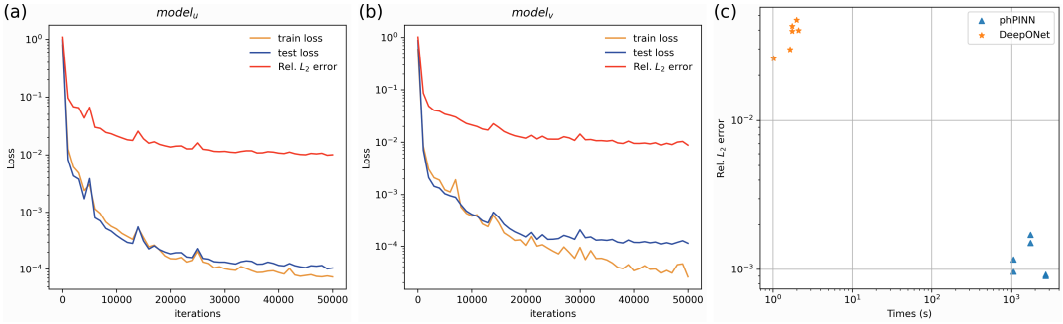

Fig. 9. (a) convergence behavior of the DeepONet training for *u*. (b) convergence behavior of the DeepONet training for *v*. (c) performance Comparison between DeepONet and phPINN.

4. Conclusion

This work studies the nondegenerate solitons of the CNLSE (1) by the phPINN (a fusion of parallel subnet and hard constraints in the PINN method) and DeepONet (an operator learning approach). The research results show that phPINN can successfully predict soliton molecules, symmetric double-double-hump solitons, and symmetric double-hump-single-hump solitons and demonstrates very low training loss and good consistency with analytical solutions, thereby showcasing the capability of phPINN as a forward solver. The phPINN algorithm also effectively addresses parameter estimation issues related to the CNLSE, demonstrating its ability to retrieve parameters under various noise conditions accurately. This is a first work to study the nonlinear optical problems using the neural operator, the numerical experiments show that neural networks successfully learn the mapping between initial conditions and solutions of nondegenerate solitons. We believe that neural network approaches in operator learning will play an increasingly important role in studying nonlinear localized waves.

Acknowledge. This work was supported by the National Natural Science Foundation of China (Grant Nos. 12261131495, 11975172, and 12381240286).

Conflict of interest. The authors have declared that no conflict of interest exists.

References

1. Li, N., Chen, Q., Triki, H., Liu, F., Sun, Y., Xu, S., & Zhou, Q. (2024). Bright and Dark Solitons in a (2+ 1)-Dimensional Spin-1 Bose-Einstein Condensates. *Ukrainian Journal of Physical Optics*, *25*(5), S1060-S1074.
2. Zhou, Q., Triki, H., Xu, J., Zeng, Z., Liu, W., & Biswas, A. (2022). Perturbation of chirped localized waves in a dual-power law nonlinear medium. *Chaos, Solitons & Fractals*, *160*, 112198.
3. Zhou, Q., Zhong, Y., Triki, H., Sun, Y., Xu, S., Liu, W., & Biswas, A. (2022). Chirped bright and kink solitons in nonlinear optical fibers with weak nonlocality and cubic-quantic-septic nonlinearity. *Chinese Physics Letters*, *39*(4), 044202.
4. Zhou, Q. (2022). Influence of parameters of optical fibers on optical soliton interactions. *Chinese Physics Letters*, *39*(1), 010501.
5. Zhong, Y., Triki, H., & Zhou, Q. (2023). Analytical and numerical study of chirped optical solitons in a spatially inhomogeneous polynomial law fiber with parity-time symmetry potential. *Communications in Theoretical Physics*, *75*(2), 025003.
6. Zhong, Y., Triki, H., & Zhou, Q. (2024). Bright and kink solitons of time-modulated cubic–quintic–septic–nonic nonlinear Schrödinger equation under space-time rotated PT-symmetric potentials. *Nonlinear Dynamics*, *112*(2), 1349-1364.

7. Ding, C. C., Zhou, Q., Triki, H., & Hu, Z. H. (2022). Interaction dynamics of optical dark bound solitons for a defocusing Lakshmanan-Porsezian-Daniel equation. *Optics Express*, *30*(22), 40712-40727.
8. Ding, C., Zhou, Q., Xu, S., Triki, H., Mirzazadeh, M., & Liu, W. (2023). Nonautonomous breather and rogue wave in spinor Bose–Einstein condensates with space-time modulated potentials. *Chinese Physics Letters*, *40*(4), 040501.
9. Liu, F. Y., Triki, H., & Zhou, Q. (2024). Optical nondegenerate solitons in a birefringent fiber with a 35 degree elliptical angle. *Optics Express*, *32*(2), 2746-2765.
10. Chen, S., Ye, Y., Soto-Crespo, J. M., Grelu, P., & Baronio, F. (2018). Peregrine solitons beyond the threefold limit and their two-soliton interactions. *Physical Review Letters*, *121*(10), 104101.
11. Kaup, D. J., & Malomed, B. A. (1993). Soliton trapping and daughter waves in the Manakov model. *Physical Review A*, *48*(1), 599.
12. Akramov, M., Sabirov, K., Matrasulov, D., Susanto, H., Usanov, S., & Karpova, O. (2022). Nonlocal nonlinear Schrödinger equation on metric graphs: A model for generation and transport of parity-time-symmetric nonlocal solitons in networks. *Physical Review E*, *105*(5), 054205.
13. Menyuk, C. R. (2024). Solitons in birefringent optical fibers and polarization mode dispersion. *Optics Communications*, *550*, 129841.
14. Wang, X. M., Zhang, L. L., & Hu, X. X. (2020). Various types of vector solitons for the coupled nonlinear Schrödinger equations in the asymmetric fiber couplers. *Optik*, *219*, 164989.
15. Zhong, H., Tian, B., Jiang, Y., Li, M., Wang, P., & Liu, W. J. (2013). All-optical soliton switching for the asymmetric fiber couplers. *The European Physical Journal D*, *67*, 1-15.
16. Geng, K. L., Mou, D. S., & Dai, C. Q. (2023). Nondegenerate solitons of 2-coupled mixed derivative nonlinear Schrödinger equations. *Nonlinear Dynamics*, *111*(1), 603-617.
17. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422-440.
18. Wang, S., Yu, X., & Perdikaris, P. (2022). When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, *449*, 110768.
19. Sirignano, J., & Spiliopoulos, K. (2018). DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, *375*, 1339-1364.
20. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686-707.
21. Wen, G., Li, Z., Azizzadenesheli, K., Anandkumar, A., & Benson, S. M. (2022). U-FNO—An enhanced Fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources*, *163*, 104180.
22. Wang, X., Wu, Z., Song, J., Han, W., & Yan, Z. (2024). Data-driven soliton solutions and parameters discovery of the coupled nonlinear wave equations via a deep learning method. *Chaos, Solitons & Fractals*, *180*, 114509.
23. Jiang, J. H., Si, Z. Z., Dai, C. Q., & Wu, B. (2024). Prediction of multipole vector solitons and model parameters for coupled saturable nonlinear Schrödinger equations. *Chaos, Solitons & Fractals*, *180*, 114581.
24. Miao, Z., & Chen, Y. (2023). VC-PINN: Variable coefficient physics-informed neural network for forward and inverse problems of PDEs with variable coefficient. *Physica D: Nonlinear Phenomena*, *456*, 133945.
25. Lin, S., & Chen, Y. (2022). A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions. *Journal of Computational Physics*, *457*, 111053.
26. Lu, L., Pestourie, R., Yao, W., Wang, Z., Verdugo, F., & Johnson, S. G. (2021). Physics-informed neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing*, *43*(6), B1105-B1132.
27. Pang, G., Lu, L., & Karniadakis, G. E. (2019). fPINNs: Fractional physics-informed neural networks. *SIAM Journal on Scientific Computing*, *41*(4), A2603-A2626.
28. Xu, S. Y., Zhou, Q., & Liu, W. (2023). Prediction of soliton evolution and equation parameters for NLS–MB equation based on the phPINN algorithm. *Nonlinear Dynamics*, *111*(19), 18401-18417.
29. Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, *3*(3), 218-229.
30. Lin, C., Li, Z., Lu, L., Cai, S., Maxey, M., & Karniadakis, G. E. (2021). Operator learning for predicting multiscale bubble growth dynamics. *The Journal of Chemical Physics*, *154*(10).
31. Di Leoni, P. C., Lu, L., Meneveau, C., Karniadakis, G. E., & Zaki, T. A. (2023). Neural operator prediction of linear instability waves in high-speed boundary layers. *Journal of Computational Physics*, *474*, 111793.
32. Osorio, J. D., Wang, Z., Karniadakis, G., Cai, S., Chryssostomidis, C., Panwar, M., & Hovsapian, R. (2022). Forecasting solar-thermal systems performance under transient operation using a data-driven machine learning approach based on the deep operator network architecture. *Energy Conversion and Management*, *252*, 115063.
33. Mao, S., Dong, R., Lu, L., Yi, K. M., Wang, S., & Perdikaris, P. (2023). Ppdonet: Deep operator networks for fast prediction of steady-state solutions in disk–planet systems. *The Astrophysical Journal Letters*, *950*(2), L12.

34. Goswami, S., Yin, M., Yu, Y., & Karniadakis, G. E. (2022). A physics-informed variational DeepONet for predicting crack path in quasi-brittle materials. *Computer Methods in Applied Mechanics and Engineering*, *391*, 114587.
35. Zhu, M., Zhang, H., Jiao, A., Karniadakis, G. E., & Lu, L. (2023). Reliable extrapolation of deep neural operators informed by physics or sparse observations. *Computer Methods in Applied Mechanics and Engineering*, *412*, 116064.
36. Wang, S., Wang, H., & Perdikaris, P. (2021). Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Science Advances*, *7*(40), eabi8605.
37. Jin, P., Meng, S., & Lu, L. (2022). MIONet: Learning multiple-input operators via tensor product. *SIAM Journal on Scientific Computing*, *44*(6), A3490-A3514.
38. Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, *63*(1), 208-228.

Su-Yong Xu, Ao-Cheng Yang, Qin Zhou. (2024). Prediction of Nondegenerate Solitons and Parameters in Nonlinear Birefringent Optical Fibers Using PhPINN and DeepONet Algorithms. *Ukrainian Journal of Physical Optics*, *25*(5), S1137 – S1150.
doi: 10.3116/16091833/Ukr.J.Phys.Opt.2024.S1137

Анотація*. У цій статті використовуються паралельні нейронні мережі з жорсткими обмеженнями, засновані на фізичних даних, для дослідження передбачення невиродженого солітону та оцінки параметрів для зв'язаного нелінійного рівняння Шредінгера. На основі наших попередніх аналітичних результатів у прямій задачі було передбачено три типи невироджених солітонів за відповідних початкових і граничних умов. У зворотній задачі, коли в якості навчального набору використовуються чисті дані, відносні похибки у передбаченні параметрів системи дисперсії групової швидкості β_2 та нелінійності Керра γ становлять менше ніж 1%. Крім того, при введенні 5% рівня шуму в навчальну множину відносні похибки для β_2 і γ залишаються нижчими ніж 3%. Крім того, ми вперше представляємо застосування Deep Operator Networks для прогнозування невироджених солітонів, зменшуючи відносну помилку L2 до 10^{-3} і досягаючи прискорення приблизно в 10^3 рази більшого порівняно з методом phPINN. Це демонструє ефективність методів навчання операторів у вирішенні нелінійних оптичних проблем.*

Ключові слова*: невироджені солітони, зв'язане нелінійне рівняння Шредінгера, phPINN, DeepONet*