Physics-Informed Neural Network Method for Predicting Soliton Dynamics Supported by Complex Parity-Time Symmetric Potentials

Xi-Meng Liu(刘希萌), Zhi-Yang Zhang(张之阳), and Wen-Jun Liu(刘文军)*

School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

(Received 5 May 2023; accepted manuscript online 26 May 2023)

We examine the deep learning technique referred to as the physics-informed neural network method for approximating the nonlinear Schrödinger equation under considered parity-time symmetric potentials and for obtaining multifarious soliton solutions. Neural networks to found principally physical information are adopted to figure out the solution to the examined nonlinear partial differential equation and to generate six different types of soliton solutions, which are basic, dipole, tripole, quadruple, pentapole, and sextupole solitons we consider. We make comparisons between the predicted and actual soliton solutions to see whether deep learning is capable of seeking the solution to the partial differential equation described before. We may assess whether physics-informed neural network is capable of effectively providing approximate soliton solutions through the evaluation of squared error between the predicted and numerical results. Moreover, we scrutinize how different activation mechanisms and network architectures impact the capability of selected deep learning technique works. Through the findings we can prove that the neural networks model we established can be utilized to accurately and effectively approximate the nonlinear Schrödinger equation under consideration and to predict the dynamics of soliton solution.

DOI: 10.1088/0256-307X/40/7/070501

Characteristics and applications of solitons have been extensively investigated in many disciplines of optics over the last decades. Especially, a great deal of attention has been paid to utilizing rich advancements of solitons in nonlinear fiber optics and telecommunication.[1–3] In mathematics, the dynamical behavior of soliton pulse propagation is illustrated by the nonlinear Schrödinger (NLS) equation, which can form steady solitons with appropriate dispersion and nonlinear parameters.[4,5]

Since spatial fractional quantum mechanics were first presented in optics and optically implemented based on transversal optical dynamics in aspheric light cavities,[6] it is naturally applied to the nonlinear optics field and modeled to utilize nonlinear fractional Schrödinger equations (NLFSEs). The latest research in the field indicates that NLFSEs generate all sorts of fractional optical solitons, e.g., vortex solitons,[7] one-dimensional gap solitons,[8] symmetric, anti symmetric solitons,[9] and two-dimensional optical solitons.[10]

Meanwhile, research about the parity-time (PT) symmetric potential introduced in quantum mechanics has sprung up. Indeed, Bender *et al.* identified the existence of a non-Hermitian Hamiltonian having a positive and real spectrum and investigated the characters of PT symmetry.[11] The conceptual theory was subsequently expanded to optical field,[12] where the external potential of the PT symmetric optical system fulfills the requirement $V(x) = V^*(x)$, which has an odd-even operator P conversion $P\psi(r,x) = \psi(-r,x)$ and a time reversal operator T conversion $T\psi(r,x) = \psi^*(-r,x)$. In particular, the dynamics of stable optical solitons supported by systems with parity-time symmetry have been explored in multifarious experiments and theoretical models, such as discrete solitons[13] and multipole solitons.[14]

With the explosive growth of computing power contemporarily, machine learning and deep learning techniques are playing vital roles in various fields of nonlinear dynamics.[15–17] For instance, in 2017, an innovative deep neural network framework was proposed by Raissi *et al.*,[18] which effectively integrated the laws of physical information governed by partial differential equations (PDEs) with deep learning. The essential principle of the physics-informed neural network (PINN) algorithm is an approximation technique to solve PDEs by converting the direct methods into loss function optimization.[19] Owing to the excellent performance, the PINN algorithm has been highly applied to solve heterogeneous NLS equations and their generalizations.[20,21] However, the PINN method has not been used to solve the saturable NLSEs and NLFSEs under considered PT symmetric potential and to make comprehensive studies of the fundamental, dipole, and multipole soliton solutions.

Firstly, we consider the (1+1)-dimensional NLS equation

$$i\frac{\partial\psi}{\partial z} + \frac{\partial^2\psi}{\partial\xi^2} + U(\xi)\psi + \sigma|\psi|^2\psi = 0, \tag{1}$$

which describes the beam propagation with PT symmetric potential. Secondly, we consider the NLFSE[22]

$$i\frac{\partial\psi}{\partial z} - \Big(-\frac{\partial^2}{\partial\xi^2}\Big)^{\alpha/2}\psi + U(\xi)\psi + \frac{\sigma|\psi|^2\psi}{1+S|\psi|^2} = 0, \tag{2}$$

which controls the transmission dynamics of light waves with PT symmetric potential and self-defocusing saturable

*Corresponding author. Email: jungliu@bupt.edu.cn
© 2023 Chinese Physical Society and IOP Publishing Ltd

nonlinearity in paraxial area. In the above two equations, $\psi(z,\xi)$ denotes the complex amplitude of the electromagnetic field and $U(\xi) = V(\xi) + iW(\xi)$ represents the normalized PT symmetric complex potential that consists of an odd gain loss and an even refractive index modulation. The coefficient $\sigma = 1/-1$ separately indicates the focusing and defocusing nonlinearity, and S describes the system's saturable nonlinear intensity. In the present work, we consider using the PINN algorithm to predict the fundamental, dipole and multipole soliton solutions to Eqs. (1) and (2), and initially deliberate the impact of different activation functions and neural network structures on performance of neural networks based on physical information.

The remainder of the material is organized logically as follows: We first make a concise introduction on the methodologies of standard PINN algorithm and present how the deep learning approach is applied for solving the considered NLSE and NLFSE. Then, we apply the PINN deep learning method to forecast fundamental and dipole solitons for the NLSE and multipole solitons for the the NLFSE, compare the result with numerical solutions, and give the error produced by neural networks. In addition, we explore the influence of the factors of the PINN algorithm on solution performance. Finally, we summarize this work and provide an outlook on the developments.

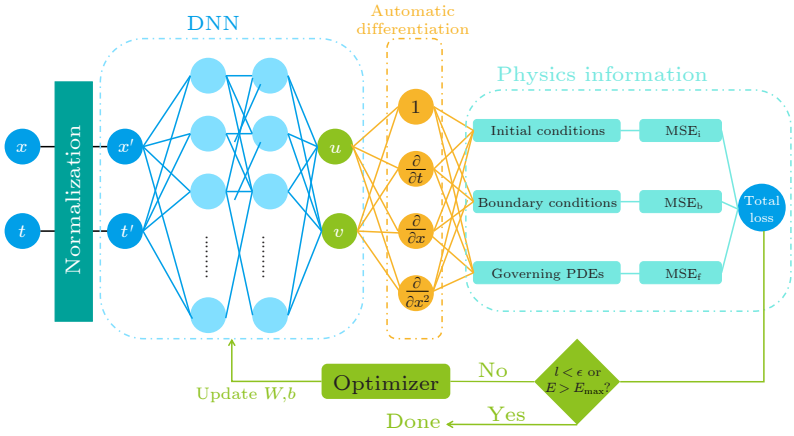

Fig. 1. A typical PINN framework. E: the number of epochs used to train a neural network, $E_{\max}$: the maximum of E, ϵ: threshold of training loss, l: training loss of the neural network.

To explain the concept and working principle of the PINN algorithm, we consider a partial differential equation in form of

$$\begin{aligned} &\lambda\frac{\partial^n u(x,t)}{\partial x^n} + f(x,t) = 0,\\ &u(0,t) = \phi(0,t),\\ &u(x,0) = g(x,0), \end{aligned} \tag{3}$$

where t and x denote the variables of solution $u(x,t)$; λ is a known parameter within the equation; $\phi(0,t)$ and $g(x,0)$ represent the given Dirichlet boundary and the initial condition, respectively. An emblematic PINN-frame structure to resolve Eq. (3) is presented in Fig. 1.

In the first stage of PINN algorithm flow, the training data of two independent variables (x,t) is processed and used in the input layer of the neural network. The input data comprise three elements: (1) the initial condition points $(x_{\text{initial}}, 0)$ sampled at $t = 0$; (2) the boundary condition points $(x_{\text{boudary}}, t_{\text{boudary}})$ sampled at the boundary condition; (3) the allocation equation points (x,t) sampled in the designated area. Subsequently, a completely connected dense neural network (DNN) including the parameter weight W and bias b is exploited to receive input points (x_i, t_i) and to obtain the corresponding predictive data $u_{\text{pred}}(x_i; t_i; *(W,b))$ that is approximately consistent with the precise $u(x,t)$. A DNN can provide powerful technique support called automatic differentiation [21] for effectively and accurately evaluating the partial derivatives of $u_{\text{pred}}(x_i; t_i; *(W,b))$ in regard to x_i and t_i, which are adopted to compute the total mean square errors $\text{MSE}_{\text{total}}$. Ultimately the network parameters W and b will be constantly updated and $u_{\text{pred}}(x_i; t_i; *(W,b))$ can convergence to the exact solution $u(x,t)$ using the minimization of the loss function $\text{MSE}_{\text{total}}$:

$$\text{MSE}_{\text{total}} = \text{MSE}_{\text{i}} + \text{MSE}_{\text{b}} + \text{MSE}_{\text{f}}, \tag{4}$$

with

$$\begin{aligned} \text{MSE}_{\text{i}} &= \frac{1}{N_{\text{i}}}\sum_{i=1}^{N_{\text{i}}}\left|u_{\text{pred}}(x_i^{\text{initial}}, 0; \theta) - g(x_i^{\text{initial}}, 0)\right|^2,\\ \text{MSE}_{\text{b}} &= \frac{1}{N_{\text{b}}}\sum_{i=1}^{N_{\text{b}}}\left|u_{\text{pred}}(0, t_i^{\text{boundary}}; \theta) - \phi(0, t_i^{\text{boundary}})\right|^2,\\ \text{MSE}_{\text{f}} &= \frac{1}{N_{\text{f}}}\sum_{i=1}^{N_{\text{f}}}\left|\lambda\frac{\partial^n u_{\text{pred}}(x_i^{\text{f}}, t_i^{\text{f}}; \theta)}{\partial x^n} + f(x_i^{\text{f}}, t_i^{\text{f}})\right|^2, \end{aligned} \tag{5}$$

where MSE_{i}, MSE_{b}, and MSE_{f} represent the loss of initial points, boundary points, and internal allocation points, respectively; while N_{i}, N_{b}, and N_{f} indicate the amounts of sampling points in the corresponding domain. Due to the characteristics of automatic differentiation and continuous prediction, PINN algorithm is mesh-free so that it is not restricted to any rigid point and is capable of predicting the $u_{\text{pred}}(x,t)$ corresponding to arbitrary points, hence sufficiently preventing the truncation and discretization mistakes inherent in conventional numerical approaches.

In order to attain the soliton solution of Eq. (1) utilizing the PINN approach mentioned above, we take the

NLS equation with PT symmetric potential under the initial and boundary conditions as follows:

$$\begin{aligned} &i\psi_z = -\psi_{\xi\xi} - U(\xi)\psi - \sigma|\psi|^2\psi, \quad \xi \in (-L, L), \\ &\psi(\xi, 0) = \psi_0(\xi), \\ &\psi(-L, z) = \psi(L, z). \end{aligned} \tag{6}$$

Because the solution $\psi(\xi, z)$ denotes a complex field, to apply the PINN method we must change the solution into the form $\psi(\xi, z) = u(\xi, z) + iv(\xi, z)$. Now, the PINN algorithm can be transformed into a complex-valued mode as $f(\xi, z) = if_u(\xi, z) - f_v(\xi, z)$. Therefore, the explicit form for the NLSE case is given as follows:

$$\begin{aligned} &f(\xi, z) = i\psi_z + \psi_{\xi\xi} + [V(\xi) + iW(\xi)]\psi + \sigma|\psi|^2\psi, \\ &f_u(\xi, z) = u_z + v_{\xi\xi} + V(\xi)v + W(\xi)u + \sigma(u^2 + v^2)v, \\ &f_v(\xi, z) = v_z - u_{\xi\xi} - V(\xi)u + W(\xi)v - \sigma(u^2 + v^2)u. \end{aligned} \tag{7}$$

By means of the proposed method, the solution $\psi(\xi, z)$ can be approximated with the PINN method. As mentioned above, all the parameters of the complex-valued neural network can be trained by minimization of loss function $\text{Loss}_{\text{train}}$ like Eq. (4):

$$\text{Loss}_{\text{train}} = \text{Loss}_{\text{i}} + \text{Loss}_{\text{b}} + \text{Loss}_{\text{f}}, \tag{8}$$

which consists of the three mean squared errors as outlined below:

$$\begin{aligned} \text{Loss}_{\text{i}} &= \frac{1}{N_{\text{i}}} \sum_{j=1}^{N_{\text{i}}} \left\{ \left| u_{\text{pred}}(\xi_{\text{i}}^j, 0) - u_0^j \right|^2 + \left| v_{\text{pred}}(\xi_{\text{i}}^j, 0) - v_0^j \right|^2 \right\}, \\ \text{Loss}_{\text{b}} &= \frac{1}{N_{\text{b}}} \sum_{j=1}^{N_{\text{b}}} \left\{ \left| u_{\text{pred}}(-L, z_{\text{b}}^j) - u_{\text{pred}}(L, z_{\text{b}}^j) \right|^2 \right. \\ &\quad \left. + \left| v_{\text{pred}}(-L, z_{\text{b}}^j) - v_{\text{pred}}(L, z_{\text{b}}^j) \right|^2 \right\}, \\ \text{Loss}_{\text{f}} &= \frac{1}{N_{\text{f}}} \sum_{j=1}^{N_{\text{f}}} \left\{ \left| f_u(\xi_{\text{C}}^j, z_{\text{C}}^j) \right|^2 + \left| f_v(\xi_{\text{C}}^j, z_{\text{C}}^j) \right|^2 \right\}, \end{aligned} \tag{9}$$

with $\{\xi_{\text{i}}^j, u_0^j, v_0^j\}_{j=1}^{N_{\text{i}}}$ representing the data under the initial condition, $\{z_{\text{b}}^j, u(\pm L, z_{\text{b}}^j), v(\pm L, z_{\text{b}}^j)\}_{j=1}^{N_{\text{b}}}$ presenting the data under the boundary condition, and $\{\xi_{\text{f}}^j, z_{\text{C}}^j, f_u(\xi_{\text{C}}^j, z_{\text{C}}^j), f_v(\xi_{\text{C}}^j, z_{\text{C}}^j)\}_{j=1}^{N_{\text{f}}}$ containing the collocation data on $f(\xi, z)$. Figure 2 displays the diagrammatic sketch of the PINN algorithm applied on the NLSE with PT symmetric potential. Deep neural network which possess two input neurons and two output neurons is presented in the left pane. The physical information of the neural network regarded as the loss function of the optimizer $\text{Loss}_{\text{train}}$ is shown on the right.

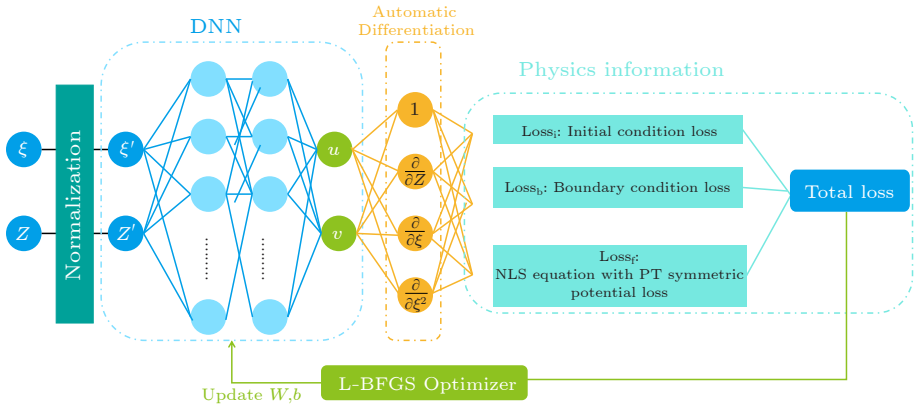

Fig. 2. Diagrammatic sketch of the PINN algorithm applied on the NLSE with PT symmetric potential.

The training dataset is composed of $N_{\text{i}} = 50$ determined initializing condition data, $N_{\text{b}} = 100$ determined boundary condition data and $N_{\text{f}} = 20000$ randomized selecting collocation points sampled by the LHS (Latin hypercube sampling) approach.[23] We establish a deep neural network with six layers, in which the first and final layers include two neurons that deal with input and output data (z, ξ) and $\psi(\xi, z)$. Every hidden layer remains a neural network that has 100 units of neurons, and hyperbolic tangent activation[24] is employed as the activation function.

Then, we focus on solutions with the properties of real physics, which indicates that the proposed solutions are stable and have real eigenvalues. To derive the optical localized mode of Eq. (2), it is assumed that the related stable solution takes the form as follows:

$$\Psi(z, \xi) = u(\xi)e^{i\beta z}, \tag{10}$$

where β denotes the propagation constant and $u(\xi) \equiv u_r + iu_i$ can be described by the equation

$$-\Big(-\frac{d^2}{d\xi}\Big)^{\alpha/2} u + V(\xi)u - \frac{|u|^2 u}{1 + S|u|^2} - \beta u = 0. \tag{11}$$

Since the model is non-integrable, we employ the power-conserving square operator method (PCSOM),[25] in which $P = \int_{-\infty}^{\infty} |u|^2 d\xi$ denotes the formulization of the power to solve Eq. (11) and to collect the multipole solitons of Eq. (1). Simultaneously, for obtaining stable data of multipole solitons in the NLFSE, we substitute two turbulent solutions $v(\xi)$ and $\vartheta(\xi)$ into the stable solutions $u(\xi)$, then Eq. (10) takes the following form:

$$\Psi(z, \xi) = e^{i\beta z}[u(\xi) + ve^{\delta z} + \vartheta^* e^{\delta^* z}]. \tag{12}$$

Substituting Eq. (12) into Eq. (2) and implementing linearization, we can achieve the characteristic equation in

the form of

$$
\begin{aligned}
-i\delta v =& -\Big(-\frac{d^2}{d\xi^2}\Big)^{\alpha/2} v + (U-\beta)v \\
&+ \frac{2\sigma|u|^2}{1+S|u|^2}v - \frac{S\sigma|u|^4}{(1+S|u|^2)^2}v \\
&+ \frac{\sigma\psi^2}{1+S|u|^2}\vartheta - \frac{S\sigma\psi^2|u|^2}{(1+S|u|^2)^2}\vartheta, \\
-i\delta\vartheta =& +\Big(-\frac{d^2}{d\xi^2}\Big)^{\alpha/2} \vartheta + (\beta-U^*)\vartheta - \frac{2\sigma|u|^2}{1+S|u|^2}\vartheta \\
&+ \frac{S\sigma|u|^4}{(1+S|u|^2)^2}\vartheta - \frac{\sigma(u^*)^2}{1+S|u|^2}v \\
&+ \frac{S\sigma(u^*)^2|u|^2}{(1+S|u|^2)^2}v. \qquad (13)
\end{aligned}
$$

The Fourier collocation method is employed to solve Eq. (13). Therefore, we can use the numerical solutions to construct datasets for the PINN method. In addition, the loss function, which should be minimized to train the neural network, has the form of

$$\text{Loss}_{\text{train}} = |\Psi_{\text{numerical}}| - |\Psi_{\text{prediction}}|, \qquad (14)$$

where $\Psi_{\text{numerical}}$ denotes numerical solution and $\Psi_{\text{prediction}}$ stands for predicted solution.

The training set comprises 2000 points sampled by the Latin method. We employ a six-layer neural network that exploits input (ξ, z) and output $\psi(\xi, z)$ handled by a pair of neurons in the first and final layers, respectively. Each of the remained four hidden layers comprises 100 neuronal units. Finally, hyperbolic tangent activation is adopted as the activation function, and the adaptive moment estimation (Adam) is applied to train the PINN model.

First, we are primarily concerned with outcomes obtained from the PINN algorithm for forecasting fundamental and dipole soliton dynamics supported by the NLSE under the considered PT symmetric potential. We settle $\sigma = 1$ for representing self-focusing propagation. We take PT symmetric Gaussian potential as our study object:

$$U(\xi) = V(\xi) + iW(\xi) = e^{-\xi^2} + iW_0\xi e^{-\xi^2}, \qquad (15)$$

where W_0 represents the intensity of the imaginary part and is assigned to the value of 0.1. Since the fundamental and dipole solitons have been worked thoroughly, we go directly to the PINN results. In order to confirm the effect of our method, the solutions computed by the PINN algorithm are compared with the solutions obtained by the numerical method. For generating the latter data, we apply a split-step Fourier transformation method and make a numerical simulation of the NLSE with the PT symmetric potential.

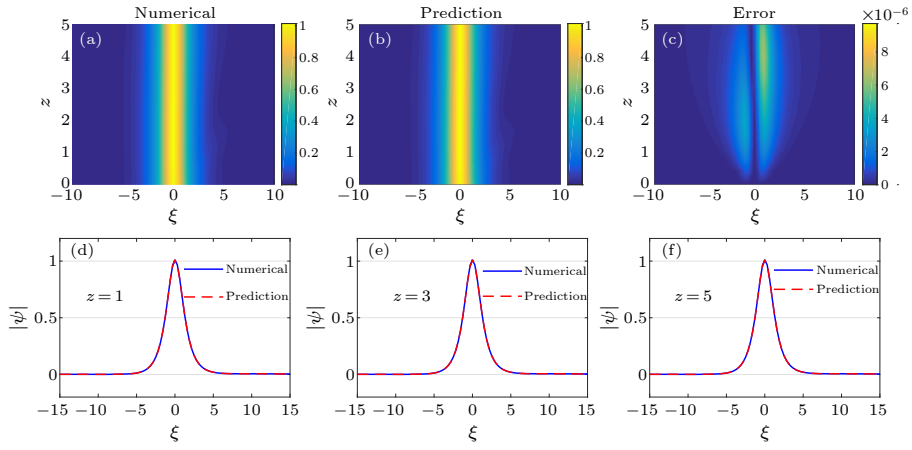

Fig. 3. PINN outcome of finding the approximate fundamental soliton solution to the nonlinear Schrödinger problem under Gaussian potential. (a) Numerical solution of the fundamental soliton. (b) Predicted dynamic behavior of the fundamental soliton. (c) Plots of difference between the predicted and numerical solutions. (d)–(f) Numerical and predicted solutions for (d) $z = 1$, (e) $z = 3$, and (f) $z = 5$.

Taking the Gaussian profile as the initial condition, we can use a neural network to obtain approximative solution of the fundamental soliton to Eq. (15) under Gaussian potential (16) mentioned above. The approximative solution is presented in Fig. 3, completing the execution of the PINN training process with the PT symmetric Gaussian potential. Figures 3(a) and 3(b) dictate the numerical and predicted soliton solutions of $|\psi(\xi, z)| = \sqrt{u^2(\xi, z) + v^2(\xi, z)}$, respectively, for the NLS equation for the given configuration of Gaussian potential. We can find from Fig. 3(a) that the predicted soliton solution produced by the PINN algorithm provides an accurate approximation for the numerical solution of the NLSE under the considered Gaussian potential (15). We illustrate the error value with drawing Fig. 3(c) in order to investigate the squared error value between predicted and numerical soliton solutions, and extrapolate that the error value is in the order of 10^{-6}. The corresponding relative L^2-norm errors of $u(\xi, z)$, $v(\xi, z)$, and $\psi(\xi, z)$ are 2.8856×10^{-2}, 2.1822×10^{-2}, and 3.0915×10^{-3}. These findings suggest that in consideration of the low error value, the proposed PINN method is capable of approximating the fundamental soliton of the NLSE with PT symmetric Gaussian potential. Figures 3(d)–3(f) depict the comparative analyses of the numerical and predicted soliton solutions at distinct propagation distance and reveal the good fitting capability

of the PINN algorithm.

Now, we explore dipole solitons which are also supported by the NLSE including PT symmetric Gaussian potential. W_0, denoting the imaginary part's intensity, should be kept at a value of 0.1. In this case, the Gaussian function remains a proper configuration as the initial condition for the purposes of approximating the dipole soliton. As noted earlier, we can attain the dipole soliton solution of the objective NLS equation computed with the PINN algorithm after training it with preliminary settings. Figures 4(a) and 4(b) dictate the numerical and predicted soliton solutions of $|\psi(\xi,z)| = \sqrt{u^2(\xi,z)+v^2(\xi,z)}$, respectively, for the NLS equation using the previously described configured Gaussian potential. Figure 4(c) illustrates the squared errors between the numerical and the predicted solution. It is clearly found that the errors range in 10^{-5}. Therefore, it is proven that the PINN method we have developed is highly accurate in approximating the dipole soliton solution for the nonlinear Schrödinger equation that contains a Gaussian potential with PT symmetry. Ulteriorly, in order to verify if the solution we obtain is correct, we plot the solution $|\psi(\xi,z)|$ at distinct propagation distance in Figs. 4(d)–4(f), and the results further approve that the predicted soliton solution acquired by the PINN algorithm can precisely approximate the numerical solution. The corresponding relative L^2-norm error of $u(\xi,z)$, $v(\xi,z)$, and $\psi(\xi,z)$ are 5.272×10^{-2}, 4.9895×10^{-2}, and 4.0925×10^{-3}, respectively.

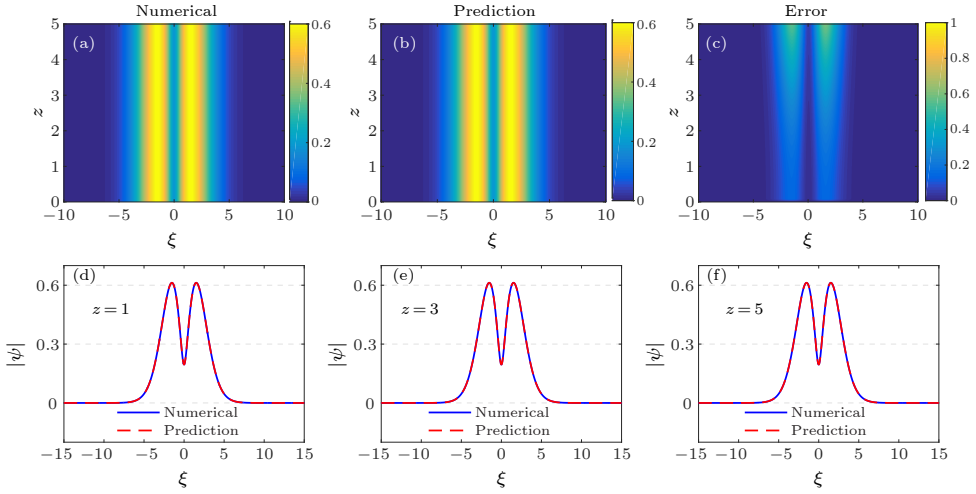

Fig. 4. PINN outcome of finding the approximate dipole soliton solution of the nonlinear Schrödinger problem under Gaussian potential. (a) Numerical solution of the dipole soliton. (b) Predicted dynamic behavior of the dipole soliton. (c) Plots of difference between the predicted and numerical solutions. (d)–(f) Numerical and predicted solutions for (d) $z=1$, (e) $z=3$, and (f) $z=5$.

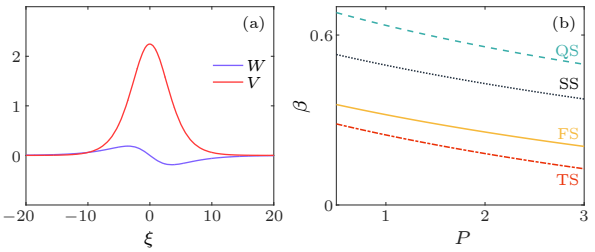

Fig. 5. (a) Real and imaginary parts (red and purple solid lines, respectively) of the PT symmetric potential with $\omega_0=1.5$ and $\kappa_0=4.0$. (b) Curves of the sextupole, pentapole, quadrupole, and tripole describing power propagation constant (black, yellow, green, and red line, respectively) with $\kappa_0=4.0$, $S=1$, and $\alpha=1.5$. For the tripole soliton with $\omega_0=1.2$ and for all other modes with $\omega_0=1.7$.

Next, we mainly analyze performance of the PINN algorithm to predict multipole soliton dynamics. First, to explain the PT symmetric potential we consider

$$U(\xi)=V(\xi)+iW(\xi)$$
$$=\omega_0^2[\mathrm{sech}(\xi/\kappa_0)]^2-i\{(\omega_0/\kappa_0)[\mathrm{sech}(\xi/\kappa_0)\tanh(\xi/\kappa_0)]\}. \quad (16)$$

We set the modulation intensity $\omega_0=1.5$ and width $\kappa_0=4.0$, thus create a structure with an odd symmetric double-peak imaginary part and an even symmetric single-peak real part, as illustrated in Fig. 5(a). Next, we attain the localized sextupole, pentapole, quadrupole, and tripole mode solitons utilizing the PCSOM, with their P–β curve depicted in Fig. 5(b). Figure 5(b) presents the power-propagation constant curve and highlights that all mode solitons exist, and thus proves the coexistence of the modes.

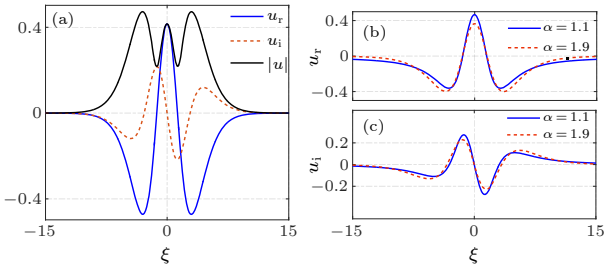

Fig. 6. (a) Real and imaginary parts (blue solid and red dotted lines, respectively) and tripole soliton modulus (black solid line) for $\alpha=2.0$. [(b), (c)] Real and imaginary part contours of the tripole soliton for $\alpha=1.1$ and $\alpha=1.9$, respectively

First, we investigate the solitons' presence. In litera-

ture, researchers have suggested that, among the numerous elements influencing the presence of solitons, the fractional diffraction effect α and the modulation intensity ω_0 have the most significant impacts. Figure 6(a) illustrates the tripole soliton's contour that comprises the odd and even symmetries of the imaginary and real parts, respectively. Figures 6(c) and 6(d) infer that by increasing the fractional diffraction effect, the outermost portions' peak amplitudes of the triple soliton's real and imaginary contours increase. Additionally, the remained locations' peak amplitudes decline while the breadth of the triple soliton increases.

Figure 7(b) reveals that the tripole soliton's predicted dynamic behavior resembles the numerical solution illustrated in Fig. 7(a). Hence, Fig. 7(c) plots the $|\Psi_{\text{numerical}}| - |\Psi_{\text{prediction}}|$, i.e., the predicted solution minus the numerical solution, which is in order of 10^{-3}. The plot demonstrates minor deviations, while the majority of the deviations are centered at the tripole soliton's peak. Figures 7(d)–7(f) compare the predicted and the calculated soliton solution at various propagation distances, e.g., $z = 1$, $z = 3$, $z = 5$, revealing that the predicted solitons fit well the numerical solutions at different propagation distances. This proves the PINN's ability to solve the nonlinear fractional Schrödinger equation including parity-time symmetric potential and saturable nonlinearity.

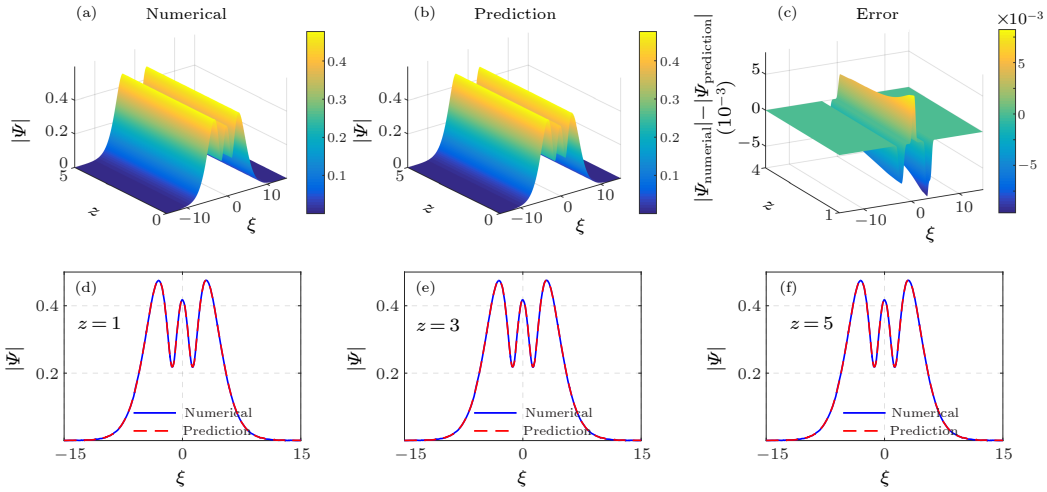

Fig. 7. PINN outcome in approximating the tripole soliton solution of the NLFSE problem. (a) Numerical solution of the tripole soliton. (b) Predicted dynamic behavior of the triple soliton. (c) The predicted solution minus the numerical solution. (d)–(f) Numerical and predicted solutions for (d) $z = 1$, (e) $z = 3$, and (f) $z = 5$.

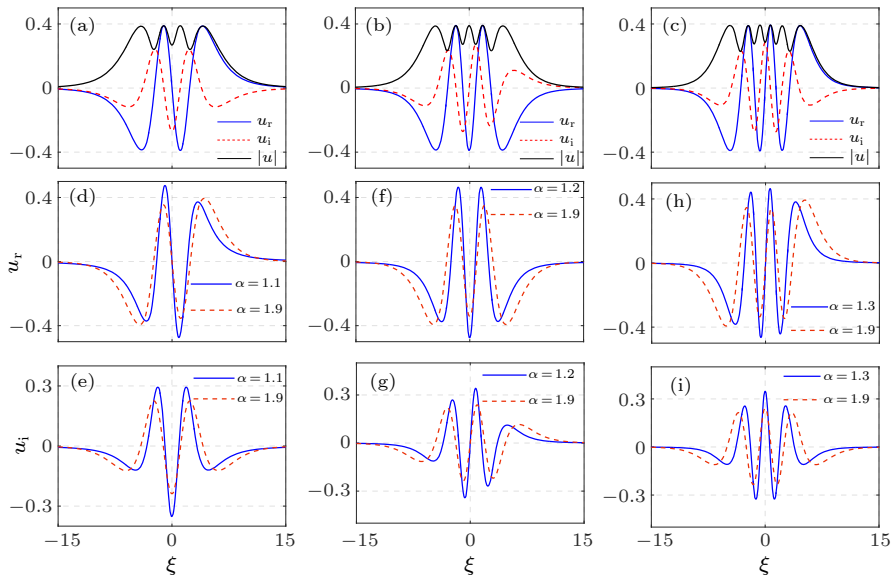

Fig. 8. Real and imaginary part distributions and modulus (blue, red, and black lines, respectively) of (a) quadrupole, (b) pentapole, and (c) sextupole solitons. Real [(d)–(h)] and imaginary [(e)–(i)] parts of [(d), (e)] quadrupole, [(f), (g)] pentapole, and [(h), (i)] sextupole solitons.

On foundation of built tripole solitons, we create sextuple, pentapole, and quadrupole solitons that have complex structures. Multipole localized modes can be created using parity-time symmetric potential, saturable nonlinearity and fractional diffraction. Similar to tripole solitons, adding ω_0 and α expands the existence range of the sextuple, pentapole, and quadrupole solitons, with their existence more clearly impacted by the fractional diffraction effect than dipole and tripole solitons due to their more intricate multi-peak structures.

Figures 8(a)–8(c) illustrate the corresponding soliton contours. We investigate modification of breadth and amplitude of the multipole soliton using the fractional diffraction effects. Figures 8(d)–8(i) depict the figurations about the real and imaginary parts of all the three multipole localize modes, revealing that these retain the tripole soliton characteristics when the fractional diffraction effect is added. Indeed, the increase in the fractional diffraction effect can result in the increment of the peak amplitudes at the furthermost profiles of imaginary and real parts for the sextupole, pentapole, and quadrupole solitons. Moreover, at other positions, the peak amplitudes decrease as the multipole soliton width increases. The experimental results suggest that the fractional diffraction effect enhances the diffraction and expansion of the multipole soliton breadth.

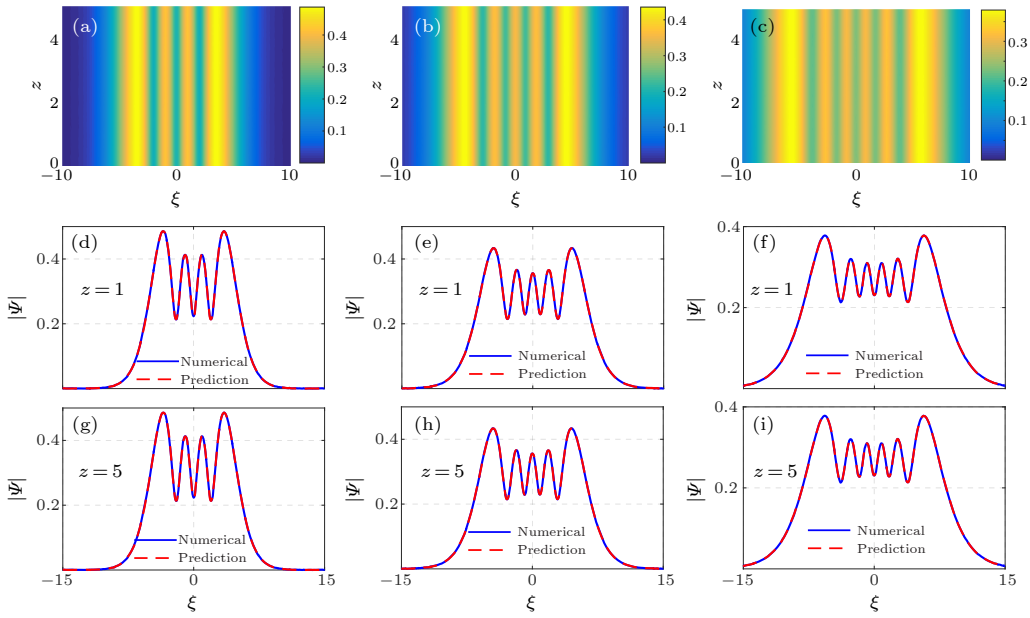

Fig. 9. Predicted dynamics of (a) quadrupole, (b) pentapole, and (c) sextupole solitons, and (d)–(i) comparative plots of numerical and predicted solutions (blue and red lines, respectively) at various evolutionary distances.

Table 1. Errors of u, v, and ψ of the PINN algorithm when approximated using different activation functions.

Solitons	Error	ReLU	sigmoid	sech	tanh
Fundamental	u	8.2281×10^{-1}	1.6449×10^{-2}	1.6481×10^{-2}	2.8856×10^{-2}
	v	7.8415×10^{-1}	2.1915×10^{-2}	1.8939×10^{-2}	2.1882×10^{-2}
	ψ	6.2519×10^{-1}	1.0385×10^{-2}	5.8659×10^{-3}	3.0915×10^{-3}
Dipole	u	1.9093	4.7094×10^{-2}	2.9019×10^{-2}	5.272×10^{-2}
	v	1.0246	4.3379×10^{-2}	2.9961×10^{-2}	4.9895×10^{-2}
	ψ	6.0433×10^{-1}	1.7092×10^{-2}	5.2880×10^{-3}	4.0925×10^{-3}

We accurately forecast the dynamics of sextuple, pentapole, and quadrupole solitons utilizing the numerical results as the training dataset of the PINN algorithm. Additionally, we compare the analytical and the projected solution's evolutionary distance maps (blue and red lines, respectively), as presented in Figs. 9(a)–9(c). It is obvious that the numerical results are in agreement with the forecast solution of multipole solitons. Thus, the PINN algorithm successfully solves the issue of conventional numerical techniques requiring extensive expertise to change parameters.

Finally, we investigate the factors affecting performance of the PINN algorithm. Because the multipole model is non-integral and the prediction results completely depend on the dataset generated, we just investigate the fundamental and dipole soliton cases. Activation functions are mathematically defined as follows:

$$\begin{aligned} Z_j &= \mathrm{ReLU}(M) = \max(0, M), \\ Z_j &= \mathrm{sigmoid}(M) = \frac{1}{1+e^{-M}}, \\ Z_j &= \mathrm{sech}M, \\ Z_j &= \tanh M, \end{aligned} \tag{17}$$

where $M = \omega_j \cdot Z_{j-1} + b_j$. Table 1 reports the overall outcome. To enhance the comparison process, we consider

the error value approximating the real part (u), imaginary part (v) and solution ψ for fundamental, dipole solitons and activation functions. The results reported infer that the PINN algorithm using ReLU as the activation function produces significant error when making the approximation of u, v, and ψ, compared to the competitor activation functions, due to the ReLU function's piecewise linearity. Regarding the sigmoid activation function, the performance is slightly mediocre. When challenging the PINN algorithm with the sech and tanh functions, the error is low when using tanh while approximating solution ψ. However, when approximating the real (u) and imaginary (v) parts, the PINN algorithm using the sech function presents high approximation accuracy for fundamental and dipole solitons considered. It should be noted that compared to other competitor activation functions, training the PINN algorithm using sech is more time-consuming.

A DNN involves many factors, such as bias and weight vectors, randomly fluctuating to minimize the target loss function. Therefore, DNN's construction impacts PINN accuracy, with the network's breadth (the number of concealed layers) and depth (the number of units per layer) being two hyper parameters that define the DNN structure. We examine the effects of these hyper parameters for the fundamental and dipole cases, as listed in Tables 2 and 3, by setting tanh as the activation function. Table 2 lists the interplay between the PINN error and the quantity of hidden layers (1 to 4), while the number of units is 100. The corresponding results highlight that PINN performance using one hidden layer is significantly inferior to employing multiple hidden layers. Furthermore, we find that a four-layer model efficiently performs the NLSE of the fundamental and dipole solitons considered. Consequently, we set the number of hidden layers in the present study to be 4.

Table 2. Errors of u, v, and ψ of the PINN algorithm when approximated using different number of hidden layers.

Solitons	Error	1	2	3	4
Fundamental	u	4.0164×10^{-2}	2.3489×10^{-2}	2.0425×10^{-2}	2.8856×10^{-2}
	v	6.4169×10^{-2}	2.1864×10^{-2}	3.3839×10^{-2}	2.1882×10^{-2}
	ψ	2.2559×10^{-2}	5.5918×10^{-3}	5.8417×10^{-3}	3.0915×10^{-3}
Dipole	u	7.2177×10^{-1}	3.8171×10^{-2}	4.9002×10^{-2}	5.272×10^{-2}
	v	9.1851×10^{-1}	3.7245×10^{-2}	4.7382×10^{-2}	4.9895×10^{-2}
	ψ	1.6633×10^{-1}	7.6633×10^{-3}	7.0433×10^{-3}	4.0925×10^{-3}

Table 3. Errors of u, v, and ψ of the PINN algorithm approximated with various neuron sizes per hidden layer.

Solitons	Error	60	70	80	90	100
Fundamental	u	1.8945×10^{-2}	1.9176×10^{-2}	1.8510×10^{-2}	2.0515×10^{-2}	2.8856×10^{-2}
	v	2.8285×10^{-2}	2.3018×10^{-2}	2.1854×10^{-2}	2.5669×10^{-2}	2.1882×10^{-2}
	ψ	4.0088×10^{-3}	4.6188×10^{-3}	4.9689×10^{-3}	4.0189×10^{-3}	3.0915×10^{-3}
Dipole	u	4.8924×10^{-2}	4.7381×10^{-2}	5.4708×10^{-2}	5.1246×10^{-2}	5.272×10^{-2}
	v	4.7821×10^{-2}	4.2622×10^{-2}	5.7069×10^{-2}	4.5090×10^{-2}	4.9895×10^{-2}
	ψ	3.8829×10^{-3}	3.8375×10^{-3}	4.3917×10^{-3}	5.6393×10^{-3}	4.0925×10^{-3}

Next, we investigate how the quantity of neurons in the hidden levels influences the PINN effectiveness. We consider the PINN algorithm using the sech activation function and the four-layer model for this investigation. Table 3 lists the results when changing the number of neurons from 60 to 100, revealing that for a 60-neuron PINN algorithm, the error is substantial. Furthermore, as the number of neurons increases, the bias and weight matrices increase, forcing the model to optimize more parameters, which decreases the error and potentially causes fluctuations. Training the PINN algorithm using 100 neurons, the NLSE solution error is low. Therefore, each concealed layer has 100 neurons.

In summary, we investigate the prediction of a family of fundamental, dipole, and multipole solitons utilizing a neural network based on physical information and analyze it in detail. In brief, on the one hand, we study the nonlinear Schrödinger equation supported by PT symmetric Gaussian potentials carefully, and approximate triumphantly the fundamental and dipole soliton solutions of the nonlinear Schrödinger equation in assistance of the PINN algorithm. On the other hand, an NLFSE incorporating fractional diffraction, PT symmetric potential, and saturable nonlinearity is presented. The fractional diffraction effect α and the modulation strength ω_0 both influence the multipole soliton's existence domain. When the soliton power is preserved, an increase of α can cause the soliton breadth to increase, and can make the peak amplitudes of multipole soliton at the furthermost portion of the real and imaginary parts increase and at other locations decrease. We then apply the PINN algorithm on the NLFSE to predict multipole soliton dynamics and achieve good results. The previous numerical techniques' need for extensive knowledge to alter parameters is efficiently overcome by the PINN method.

We also evaluate the impact of the network structure and activation function on PINN's performance. The PINN algorithm using the sech and tanh as activation functions approximates soliton solution more accurately than that using Relu and sigmoid. Additionally, we investigate how PINN's breadth and depth affect efficiency, and it is concluded that the number of concealed levels is 4, with 100 neurons per layer.

Future work will focus on predicting and analyzing all

sorts of soliton dynamics and advance of this area. Our neural network approach, however, lengthens the training periods and learning expenses. Hence, our technique will be tweaked to increase accuracy while maintaining learning effectiveness. In the interim, we will attempt to apply this approach to more models and increase its flexibility and generalization ability.

Another significant direction for future research involves the integration of the PINN algorithm with experimental or observational data. This integration allows researchers to leverage both the physics-based constraints encoded in the governing equations and the real-world measurements obtained from experiments or observations. By incorporating real data, the PINN algorithm can improve its predictive capabilities and can enhance the accuracy of the solutions provided. Ultimately, this integration can lead to new discoveries, improved understanding of physical systems, and the development of data-driven predictive models for a wide range of applications.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (Grant No. 12075034).

References

[1] Mihalache D 2017 *Rom. Rep. Phys.* **69** 403
[2] Zhang X M, Qin Y H, Ling L M, and Zhao L Z 2021 *Chin. Phys. Lett.* **38** 090201
[3] Yan Y Y and Liu W J 2021 *Chin. Phys. Lett.* **38** 094201
[4] Zhou Q 2022 *Chin. Phys. Lett.* **39** 010501
[5] Zhou Q, Zhong Y, Triki H, Sun Y Z, Xu S L, Liu W J, and Biswas A 2022 *Chin. Phys. Lett.* **39** 044202
[6] Longhi S 2015 *Opt. Lett.* **40** 1117
[7] Dong L W and Huang C M 2019 *Nonlinear Dyn.* **98** 1019
[8] Zeng L W and Zeng J H 2019 *Nonlinear Dyn.* **98** 985
[9] Cao Q H and Dai C Q 2021 *Chin. Phys. Lett.* **38** 090501
[10] Li P F, Li R J, and Dai C Q 2021 *Opt. Express* **29** 3193
[11] Bender C M and Boettcher S 1998 *Phys. Rev. Lett.* **80** 5243
[12] Klaiman S, Günther U, and Moiseyev N 2008 *Phys. Rev. Lett.* **101** 080402
[13] Driben R and Malomed B A 2011 *Opt. Lett.* **36** 4323
[14] Zhu X, Yang F, Cao S, Xie J, and He Y 2020 *Opt. Express* **28** 1631
[15] Pathak J, Lu Z, Hunt B R, Girvan M, and Ott E 2017 *Chaos* **27** 121102
[16] Amil P, Soriano M C, and Masoller C 2019 *Chaos* **29** 113111
[17] Panday A, Lee W S, Dutta S, and Jalan S 2021 *Chaos* **31** 031106
[18] Raissi M, Perdikaris P, and Karniadakis G E 2019 *J. Comput. Phys.* **378** 686
[19] Pu J C, Peng W Q, and Chen Y 2021 *Wave Motion* **107** 102823
[20] Zhou Z J and Yan Z Y 2021 *Commun. Theor. Phys.* **73** 105006
[21] Baydin A G, Pearlmutter B A, Radul A A, and Siskind J M 2018 *J. Mach. Learn. Res.* **18(153)** 1
[22] Zhao Y, Lei Y B, Xu Y X, Xu S L, Triki H, Biswas A, and Zhou Q 2022 *Chin. Phys. Lett.* **39** 034202
[23] Stein M 1987 *Technometrics* **29** 143
[24] Golev A, Iliev A, and Kyurkchiev N 2017 *Int. J. Sci. Eng. Technol.* **4** 6
[25] Li P F, Mihalache D, and Malomed B A 2018 *Phil. Trans. R. Soc. A* **376** 20170378