



Data-driven vector localized waves and parameters discovery for Manakov system using deep learning approach

Jun-Cai Pu^a, Yong Chen^{a,b,*}

^a School of Mathematical Sciences, Shanghai Key Laboratory of Pure Mathematics and Mathematical Practice, East China Normal University, Shanghai 200241, China

^b College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China

ARTICLE INFO

Article history:

Received 5 January 2022

Received in revised form 29 April 2022

Accepted 29 April 2022

Available online xxxx

Keywords:

Data-driven vector localized waves

Vector rogue waves

Parameters discovery

Manakov system

Improved PINN

ABSTRACT

An improved physics-informed neural network (IPINN) algorithm with four output functions and four physics constraints, which possesses neuron-wise locally adaptive activation function and slope recovery term, is appropriately proposed to obtain the data-driven vector localized waves, including vector solitons, breathers and rogue waves (RWs) for the Manakov system with initial and boundary conditions, as well as data-driven parameters discovery for Manakov system with unknown parameters. The data-driven vector RWs which also contain interaction waves of RWs and bright-dark solitons, interaction waves of RWs and breathers, as well as RWs evolved from bright-dark solitons are learned to verify the capability of the IPINN algorithm in training complex localized wave. In the process of parameter discovery, routine IPINN can not accurately train unknown parameters whether using clean data or noisy data. Thus we introduce parameter regularization strategy with adjustable weight coefficients into IPINN to effectively and accurately train prediction parameters, then find that once setting the appropriate weight coefficients, the training effect is better as using noisy data. Numerical results show that IPINN with parameter regularization shows superior noise immunity in parameters discovery problem.

© 2022 Elsevier Ltd.

1. Introduction

Localized waves, containing soliton, breather and rogue wave (RW), have become a significant research direction in the field of nonlinear science in recent decades [1–4]. Compared with the breathers which are localized in space or time (Akhmediev or Ma breathers [2,3]), the RWs are localized both in space and time with the heights of the RWs are two or more times those of the surrounding waves, and RW is the name given by oceanographers to isolated large amplitude waves which appears suddenly and disappear without a trace [5]. In 1964, Draper first observed the existence of RWs and called them “Freak ocean waves” [6], and found RWs widely occur both in deep ocean and in shallow water [7–9]. Hitherto, the research of RWs has gone far beyond the oceanographic background, containing other spatial-temporal continuous systems, such as water tank [10], ultra-cold bosonic gases [11], microwave transport [12], capillary waves [13], atmosphere [14] and plasma [15]. In any of these aforementioned disciplines, new discoveries which indicate that RWs may be rather universal enrich their concept and lead to progress towards a comprehensive

understanding of a phenomenon which still remains largely unexplored to obtain state-of-the-art theories on the subject.

The so-called Peregrine soliton provided a formal mathematical description of RWs, this solitary wave is the solution of the $(1+1)$ -dimensional scalar nonlinear Schrödinger equation (NLS), which is localized in both coordinate systems and describes a unique RW event [16]. This solution is also unique in the mathematical sense, due to it is expressed in terms of rational functions of coordinates, which is different from most other known soliton solutions of the NLS, and the determinant representation of the N -order RW of the NLS has been directly obtained directly by a series of row operations on matrices appeared in the N -fold Darboux transformation [17]. In a variety of complex systems, such as optical fibers [18], financial systems [19] and Bose-Einstein condensates [20], several amplitudes rather than a single one need to be considered. The resulting coupled systems can describe extreme waves more accurately than the scalar NLS model [20,21]. Furthermore, for scalar systems, the velocity of the background field has no practical effect on the mode structure of RWs, since the corresponding solutions can be correlated by Galileo transform, but the relative velocity between different component fields for the coupled systems has real physical effects, and can not be eliminated by any simple transformation [22]. In recent years, some important results have been obtained in the study of solitons, breathers and RWs for local and nonlocal multicomponent coupled nonlinear systems, such as the two-component derivative

* Corresponding author at: School of Mathematical Sciences, Shanghai Key Laboratory of Pure Mathematics and Mathematical Practice, East China Normal University, Shanghai 200241, China.

E-mail address: ychen@sei.ecnu.edu.cn (Y. Chen).

NLS [23], coupled fractional NLS [24], nonlocal M -component NLS [25], $(2 + 1)$ -dimensional generalized coupled NLS with the four-wave mixing term [26] and nonautonomous $(2 + 1)$ -dimensional coupled NLS with partially nonlocal nonlinear effect and a linear potential [27], but the research of exact and numerical solutions for local and nonlocal multicomponent coupled nonlinear systems is relatively few compared with scalar nonlinear systems in general. Therefore, the extended researches on vector RWs for multi-component coupled systems are non-trivial and meaningful.

With the revolution of hardware equipment and software technology, the explosive growth of available data and the great improvement of computer operation speed promote the application of machine learning and big data analysis technology in practice [28,29]. Furthermore, the computational complexity of neural network (NN) was greatly reduced by introducing back-propagation NN algorithm and the gradient vanishing problem was solved, deep learning was born and has attracted extensive attention [30,31]. Since that, deep learning dominated a large number of landing applications, such as face recognition [32], medical imaging [33], video surveillance [34], speech recognition [35], language understanding [36] and mathematical physics [37]. Among these applications mentioned above, a new physics-informed neural network (PINN) which controlled by mathematical physical systems based on the deep learning patterns of multi-layer NNs, have been proposed and proved to be particularly suitable for dealing with both the forward problems and highly ill-posed inverse problems by obtaining the approximate solutions of governing equations and discovering parameters involved in the governing equation are inferred from the training data, and numerical results found the PINN approach only needs less data sets to complete high-dimensional network tasks well [37]. Recently, PINN has played an important role in many physical applications [38]. Afterwards, Jagtap and collaborators proposed two different kinds of adaptive activation functions, namely global adaptive activation functions and locally adaptive activation functions, to approximate smooth and discontinuous functions as well as solutions of linear and nonlinear partial differential equations by introducing a scalable parameters in the activation function and adding a slope recovery term based on activation slope to the loss function of locally adaptive activation functions, and demonstrated the locally adaptive activation functions further improve the training speed, performance and speed up the training process of NNs [39,40].

Recently, the data-drive forward and inverse problems of nonlinear partial differential equations have attracted extensive attention and sparked interest in their research by applying the PINN [37]. Due to the abundant sample space and good properties of integrable systems, it has become a research hotspots by applying PINN into the field of integrable systems, and significant numerical calculation results and some interesting unknown phenomena will be found. Chen research group is committed to the research of integrable deep learning algorithm, and obtains a series of high-quality data-driven solutions for classical integrable systems by using PINN and improved PINN (IPINN) algorithms. Specifically, data-driven solutions with abundant dynamic behaviors for some classical nonlinear evolution partial differential equations have been obtained by utilizing the PINN framework [41]. Especially, Pu et al. first recovered the solitons, breathers and rogue wave solutions of the NLS with the aid of the PINN model [42]. The data-driven rogue periodic wave of nonlinear partial differential equation has been learned by applying multi-layer PINN for the first time [43], and the PINN is applied to high-dimensional system to solve the $(N + 1)$ -dimensional initial boundary value problem with $2N + 1$ hyperplane boundaries [44]. It is worth mentioning that a two-stage PINN method which is tailored to the nature of equations by introducing features of physical systems into NNs, is used to simulate abundant localized wave solutions of integrable equations [45]. Furthermore, a PINN approach with neuron-wise locally adaptive activation function was presented to derive rational soliton solutions and rogue wave solutions of the derivative NLS in complex space, and numerical results

demonstrated the improved approach has faster convergence and better simulation effect than classical PINN method [46,47]. In addition, other scholars have also done important works on data-driven solutions and parameter discovery of other nonlinear systems, such as higher-order NLS and defocusing NLS with the time-dependent potential [48, 49]. However, these aforementioned works utilize various PINNs with up to two outputs and two coupled physical constraints, but we need the NNs with at least four outputs and four coupled physical constraints to simulate their corresponding data-driven vector localized waves and parameters discovery for multi-component coupled nonlinear systems, this is also the difficulty we need to solve when we reconstruct a new NN to solve the data-driven vector localized waves and parameters discovery of multi-component coupled systems. To our best knowledge, although two parallel PINN with two inputs and two outputs are proposed to study the dynamic process and model parameters of the vector optical solitons for the coupled NLS in birefringent fibers [50], as well as a pre-fixed multi-stage training PINN algorithm, which can improve the convergence rate and the approximation ability of the original PINN method, is introduced to obtain data-driven vector soliton solutions of coupled NLS [51], but the vector solitons, breathers and RWs for the Manakov system which also be called two-component coupled NLS have not been studied by applying the PINN with neuron-wise locally adaptive activation function, slope recovery term, four outputs and four physical constraints [21]. Therefore, we design and propose an improved PINN (IPINN) model with neuron-wise locally adaptive activation function, slope recovery term, four outputs, four constraint equations and parameter regularization strategy, then recover the vector localized waves and learn unknown parameters of Manakov system with the corresponding initial boundary value conditions for the first time in this paper.

Recently, we proposed PINN with two inputs and three outputs to solve data-driven forward-inverse problems for Yajima-Oikawa system by using parameter regularization strategy [52]. Next, we focus on the following Manakov system with unknown parameters λ_1 and λ_2 , the expression is as follows

$$\begin{cases} i q_{1t} + \lambda_1 q_{1xx} + \lambda_2 (|q_1|^2 + |q_2|^2) q_1 = 0, x \in [L_0, L_1], t \in [T_0, T_1], \\ i q_{2t} + \lambda_1 q_{2xx} + \lambda_2 (|q_1|^2 + |q_2|^2) q_2 = 0, x \in [L_0, L_1], t \in [T_0, T_1], \\ q_1(x, T_0) = q_1^0(x), q_2(x, T_0) = q_2^0(x), x \in [L_0, L_1], \\ q_1(L_0, t) = q_1^{lb}(t), q_1(L_1, t) = q_1^{ub}(t), t \in [T_0, T_1], \\ q_2(L_0, t) = q_2^{lb}(t), q_2(L_1, t) = q_2^{ub}(t), t \in [T_0, T_1], \end{cases} \quad (1)$$

where “ i ” satisfied $i^2 = -1$ is an imaginary number, the subscripts denote the partial derivatives of the complex fields $q_1(x, t)$ and $q_2(x, t)$ with respect to the space x and time t , while the L_0 and L_1 represent the lower and upper boundaries of x respectively. Similarly, T_0 and T_1 represent the initial and final times of t respectively. Moreover, the $q_1^0(x)$ ($q_2^0(x)$) represents initial value of the $q_1(x, t)$ ($q_2(x, t)$) at $t = T_0$, the $q_1^{lb}(t)$ and $q_1^{ub}(t)$ ($q_2^{lb}(t)$ and $q_2^{ub}(t)$) are the lower and upper boundaries of the $q_1(x, t)$ ($q_2(x, t)$) corresponding to $x = L_0$ and $x = L_1$ respectively. Eq. (1) applies to a Kerr medium with the electrostrictive mechanism of nonlinearity [53], as well as to randomly birefringent fiber optic transmission links [54].

In recent years, a variety of explicit vector solitons, breathers and RWs of the Manakov system have been obtained. By utilizing the RHP approach, Yang derived the exact determinant form of vector N -soliton for Manakov system with 3×3 Lax pair in detail [55]. Priya et al. presented explicit forms of general breather, Akhmediev breather, Ma soliton, and RWs of the Manakov system by means of the Hirota bilinear method [56]. The exact RW solutions, breathers, and rogue-bright-dark solutions for the Manakov system have been constructed by the Darboux transformation [57]. A novel multiparametric vector freak solutions which feature both exponential and rational dependence on coordinates for the Manakov system has been analytically

constructed and discussed, and the family of exact solutions includes known vector Peregrine (rational) solutions has also been derived [58]. Zhao and Liu obtained the dark RWs analytically for the first time in the Manakov system, and found that two RWs can appear in the temporal-spatial distribution [59]. According to the generalized Darboux transformation, Ling et al. studied the dynamics of high-order RWs in the Manakov system, and pointed out that four fundamental RWs can emerge from second-order vector RWs in the coupled system, in contrast to the high-order ones in single-component systems [22]. Wang et al. investigated the novel higher-order localized waves account for the Manakov system are investigated by using the generalized Darboux transformation, and exhibited that two dark-bright solitons together with a second-order RW of fundamental or triangular pattern as well as two breathers together with a second-order RW of fundamental or triangular pattern coexist in the second-order localized wave for the coupled system [60]. Rao et al. employed the Kadomtsev-Petviashvili hierarchy reduction method for deriving the general vector RW of the M -coupled NLS systems which contain the Manakov system in a compact form [61]. From the previous references, we find out that the exact vector RWs of the Manakov system is mainly obtained via Darboux transformation, so it is difficult to accurately solve the vector RWs of the multi-component coupled nonlinear systems, and the means are extremely scarce, especially for some non integrable coupled nonlinear systems. Therefore, we consider how to apply the deep learning NN architecture to solve multi-component coupled nonlinear systems, specifically how to recover various vector localized waves when only the initial-boundary value conditions of the multi-component coupled nonlinear systems are known. In the following, we focus on the data-driven vector localized waves which contain vector solitons, breather and RWs, as well as parameters discovery for the Manakov system with initial-boundary value conditions by utilizing IPINN pattern.

The rest of this paper is organized as follows. In Section 2, we introduce briefly discussions of the IPINN method with locally adaptive activation function for the Manakov system, where also discuss about training data, loss function, optimization method and the training environment. Moreover, the algorithm flow schematic and algorithm steps for the Manakov system based on IPINN model are exhibited in detail. In Section 3, the data-driven solitons and breather with vivid plots and dynamic behavior analyses of the Manakov system have been exhibited via IPINN model. Section 4 provides the data-driven vector RWs of the Manakov system by utilizing the IPINN approach, and related plots and dynamic analyses are revealed in detail. Section 5 presents experimental results with different trade-off norm penalty term coefficients during learning data-driven parameter discovery of Manakov system. Conclusions and discussions are given out in last section.

2. Methodology

In general, we consider the general $(1 + 1)$ -dimensional coupled nonlinear time-dependent systems with unknown parameters λ_1 and λ_2 in complex space, its general form is as shown below

$$\begin{aligned} i q_{1t} + \mathcal{N}[q_1, q_2; \lambda_1, \lambda_2] &= 0, \\ i q_{2t} + \mathcal{N}'[q_1, q_2; \lambda_1, \lambda_2] &= 0, \end{aligned} \tag{2}$$

where q_1 and q_2 are complex-valued solutions of x and t to be determined later, $\mathcal{N}[\cdot, \cdot; \lambda_1, \lambda_2]$ and $\mathcal{N}'[\cdot, \cdot; \lambda_1, \lambda_2]$ are nonlinear differential operators in space. Due to the complexity of the structure of the complex-valued solutions $q_1(x, t)$ and $q_2(x, t)$ in Eq. (2), we decompose $q_1(x, t)$ ($q_2(x, t)$) into the real part $u(x, t)$ ($m(x, t)$) and the imaginary part $v(x, t)$ ($n(x, t)$) by employing real-valued functions $u(x, t)$ and $v(x, t)$ ($m(x, t)$ and $n(x, t)$), that is $q_1(x, t) = u(x, t) + iv(x, t)$ and $q_2(x, t) = m(x, t) + in(x, t)$. After substituting it into Eq. (2), then letting the real and imaginary parts be equal to 0, we have

$$\begin{aligned} -v_t + \mathcal{N}_u[u, v, m, n; \lambda_1, \lambda_2] &= 0, & u_t + \mathcal{N}_v[u, v, m, n; \lambda_1, \lambda_2] &= 0, \\ -n_t + \mathcal{N}'_m[u, v, m, n; \lambda_1, \lambda_2] &= 0, & m_t + \mathcal{N}'_n[u, v, m, n; \lambda_1, \lambda_2] &= 0, \end{aligned} \tag{3}$$

accordingly, the $\mathcal{N}_u, \mathcal{N}_v, \mathcal{N}'_m$ and \mathcal{N}'_n are nonlinear differential operators in space. Then the physics-informed neural networks $f_u(x, t), f_v(x, t), f_m(x, t)$ and $f_n(x, t)$ can be defined as

$$\begin{aligned} f_u &:= -v_t + \mathcal{N}_u[u, v, m, n; \lambda_1, \lambda_2], & f_v &:= u_t + \mathcal{N}_v[u, v, m, n; \lambda_1, \lambda_2], \\ f_m &:= -n_t + \mathcal{N}'_m[u, v, m, n; \lambda_1, \lambda_2], & f_n &:= m_t + \mathcal{N}'_n[u, v, m, n; \lambda_1, \lambda_2]. \end{aligned} \tag{4}$$

From Ref. [46], one can know about the classical PINN method could not accurately reconstruct some solutions with complex forms in some complicated nonlinear systems, while the PINN approach with neuron-wise locally adaptive activation function and slope recovery term can improve the convergence speed and stability of the loss function in the training process, so the focus of this paper is no longer to explain the advantages of this improved PINN method, thus it is no longer compared with the traditional PINN in this paper. Therefore, considering the training accuracy, performance requirements and the structural complexity of multi-component coupled nonlinear systems, based on the aforementioned PINN method, we extend and propose an IPINN model for investigating the data-driven vector localized waves and parameters discovery of the Manakov system effectively in this paper. It changes the slope of the activation function adaptively, resulting in non-vanishing gradients and faster training of the network.

We establish an IPINN of depth D corresponding to the NN with an input layer, $D - 1$ hidden-layers and an output layer. In the d th hidden-layer, N_d number of neurons are present. Each hidden-layer of the IPINN receives an output $\mathbf{x}^{d-1} \in \mathbb{R}^{N_{d-1}}$ from the previous layer, where an affine transformation can be written as follows form

$$\mathcal{L}_d(\mathbf{x}^{d-1}) \triangleq \mathbf{W}^d \mathbf{x}^{d-1} + \mathbf{b}^d, \tag{5}$$

where the network weights $\mathbf{W}^d \in \mathbb{R}^{N_d \times N_{d-1}}$ and bias term $\mathbf{b}^d \in \mathbb{R}^{N_d}$ associated with the d th layer. Specifically, we define such neuron-wise locally adaptive activation function as

$$\sigma(n a_i^d (\mathcal{L}_d(\mathbf{x}^{d-1}))_i), d = 1, 2, \dots, D-1, i = 1, 2, \dots, N_d,$$

where σ is the activation function, and $n > 1$ is a scaling factor and $\{a_i^d\}$ are additional $\sum_{d=1}^{D-1} N_d$ parameters to be optimized. Note that, there is a critical scaling factor n_c , and the optimization algorithm will become sensitive when $n \geq n_c$ in each problem set. The neuron activation function acts as a vector activation function in each hidden layer, and each neuron has its own slope of activation function.

The IPINN with neuron-wise locally adaptive activation function can be represented as

$$\begin{aligned} q(\mathbf{x}; \bar{\Theta}) &= ((\mathcal{L}_D)_i \circ \sigma \circ n a_i^{D-1} (\mathcal{L}_{D-1})_i \cdots \sigma \circ n a_i^1 (\mathcal{L}_1)_i)(\mathbf{x}), \\ &= 1, 2, 3, 4, \end{aligned} \tag{6}$$

where \mathbf{x} and $q(\mathbf{x}; \bar{\Theta})$ represent the two inputs and four outputs in the IPINN, respectively. The set of trainable parameters $\bar{\Theta} \in \mathcal{P}$ consists of $\{\mathbf{W}^d \mathbf{b}^d\}_{d=1}^D$, unknown parameters (λ_1 and λ_2) to be learned and $\{a_i^d\}_{d=1}^{D-1}, \forall i = 1, 2, \dots, N_d, \mathcal{P}$ is the parameter space. In this paper, the initialization of scalable parameters are carried out in the case of $n a_i^d = 1, \forall n \geq 1$.

The resulting optimization algorithm will attempt to find the optimized parameters including the weights, biases and additional coefficients in the activation to minimize the new loss function defined as

$$\mathcal{L}(\bar{\Theta}) = \text{Loss} = \text{Loss}_{q_1} + \text{Loss}_{q_2} + \text{Loss}_{f_1} + \text{Loss}_{f_2} + \text{Loss}_a, \quad (7)$$

where Loss_{q_1} , Loss_{q_2} , Loss_{f_1} and Loss_{f_2} are defined as following

$$\begin{aligned} \text{Loss}_{q_1} &= \frac{1}{N_q} \left[\sum_{j=1}^{N_q} |\hat{u}(x^j, t^j) - u^j|^2 + \sum_{j=1}^{N_q} |\hat{v}(x^j, t^j) - v^j|^2 \right], \\ \text{Loss}_{q_2} &= \frac{1}{N_q} \left[\sum_{j=1}^{N_q} |\hat{m}(x^j, t^j) - m^j|^2 + \sum_{j=1}^{N_q} |\hat{n}(x^j, t^j) - n^j|^2 \right], \end{aligned} \quad (8)$$

and

$$\begin{aligned} \text{Loss}_{f_1} &= \frac{1}{N_f} \left[\sum_{l=1}^{N_f} |f_u(x_l^f, t_l^f)|^2 + \sum_{l=1}^{N_f} |f_v(x_l^f, t_l^f)|^2 \right], \\ \text{Loss}_{f_2} &= \frac{1}{N_f} \left[\sum_{l=1}^{N_f} |f_m(x_l^f, t_l^f)|^2 + \sum_{l=1}^{N_f} |f_n(x_l^f, t_l^f)|^2 \right], \end{aligned} \quad (9)$$

where $\{x^j, t^j, u^j, v^j, m^j, n^j\}_{j=1}^{N_q}$ denote the initial-boundary value inputs data on Eqs. (3) and (4). Here $\hat{u}(x^j, t^j)$, $\hat{v}(x^j, t^j)$, $\hat{m}(x^j, t^j)$ and $\hat{n}(x^j, t^j)$ represent the optimal training outputs data through the IPINN.

Furthermore, $\{x_l^f, t_l^f\}_{l=1}^{N_f}$ represent the collocation points on networks $f_u(x, t)$, $f_v(x, t)$, $f_m(x, t)$ and $f_n(x, t)$. The last slope recovery term Loss_a in the loss function (Eq. (7)) is defined as

$$\text{Loss}_a = \frac{1}{\frac{N_a}{D-1} \sum_{d=1}^{D-1} \exp\left(\frac{\sum_{i=1}^{N_d} a_i^d}{N_d}\right)}, \quad (10)$$

where $1/N_a$ is hyper-parameter for slope recovery term Loss_a , and we all take $N_a = 100$ for dominating the loss function and ensuring that the final loss value is not too large in this paper. Here, term Loss_a forces the NN to increase the activation slope value quickly, which ensures the non-vanishing gradient of the loss function and improves the network's training speed. Consequently, Loss_{q_1} and Loss_{q_2} correspond to the loss on the initial and boundary data, the Loss_{f_1} and Loss_{f_2} penalize the Manakov system not being satisfied on the collocation points, and

the Loss_a changes the topology of Loss function and improves the convergence speed and network optimization ability. Moreover, in order to better measure the training error, we introduce \mathbb{L}_2 norm error, which is defined as follows

$$\text{Error} = \frac{\sqrt{\sum_{k=1}^N |q^{\text{exact}}(\mathbf{x}_k) - q^{\text{predict}}(\mathbf{x}_k; \bar{\Theta})|^2}}{\sqrt{\sum_{k=1}^N |q^{\text{exact}}(\mathbf{x}_k)|^2}},$$

where $q^{\text{predict}}(\mathbf{x}_k; \bar{\Theta})$ and $q^{\text{exact}}(\mathbf{x}_k)$ represent the model training prediction solution and exact analytical solution at point $\mathbf{x}_k = (x_k, t_k)$, respectively.

In order to understand the IPINN approach more clearly, the IPINN algorithm flow chart of the Manakov system is shown in following Fig. 1, where one can see the NN along with the supplementary physics-informed part, and the loss function is evaluated using the contribution from the NN part as well as the residual from the governing equation given by the physics-informed part. Then, one can seek the optimal values of weights \mathbf{W} , biases \mathbf{b} , parameters λ_1 and λ_2 as well as scalable parameter a_i^d in order to minimize the loss function below certain tolerance ε until a prescribed maximum number of iterations. From Fig. 1, since the Manakov system contains two components $q_1(x, t)$ and $q_2(x, t)$, one can see that the “NN” part has four output functions $\{u, v, m, n\}$, and there are four nonlinear equation constraints in the “PDE” part, that is, in terms of the nonlinear coupled system with more components, the number of output functions and nonlinear equation constraints of the IPINN will increase multiply. Furthermore, in order to further understand the IPINN, we also showcase the corresponding procedure steps of the IPINN with adaptive activation function and slope recovery term in the following table.

- Algorithm: IPINN algorithm with adaptive activation function and slope recovery term.
- Step 1:** Specification of training set in computational domain:
Training data: $\{x^j, t^j, u^j, v^j, m^j, n^j\}_{j=1}^{N_q}$, Residual training points: $\{x_l^f, t_l^f\}_{l=1}^{N_f}$.
 - Step 2:** Construct NN $q(x; \bar{\Theta})$ with random initialization of parameters $\bar{\Theta}$.
 - Step 3:** Construct the residual NN $\{f_u, f_v, f_m, f_n\}$ by substituting surrogate $q(x; \bar{\Theta})$ into the governing equations using automatic differentiation and other arithmetic operations.

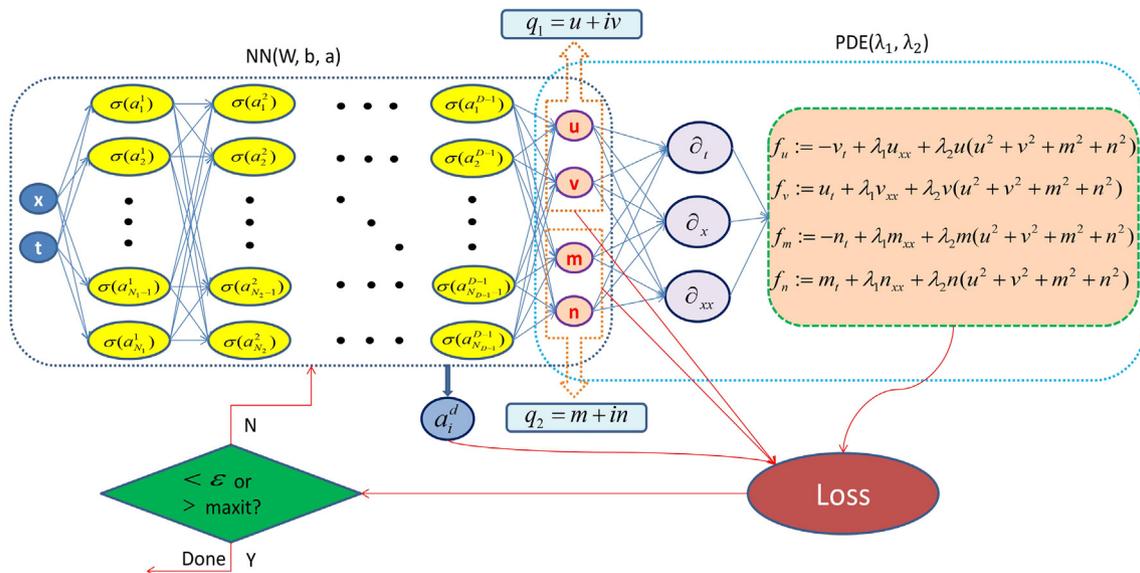


Fig. 1. Schematic of IPINN for the Manakov system with unknown parameters λ_1 and λ_2 . The left NN is the universal approximation network while the right one induced by the governing equation is the physics-informed network. The two NNs share hyper-parameters and they both contribute to the loss function.

Step 4: Specification of the loss function $\mathcal{L}(\bar{\Theta})$ that includes the slope recovery term.

Step 5: Find the best parameters $\bar{\Theta}^*$ using a suitable optimization method for minimizing the loss function $\mathcal{L}(\bar{\Theta})$ as

$$\bar{\Theta}^* = \arg \min_{\bar{\Theta} \in \mathcal{P}} \mathcal{L}(\bar{\Theta}).$$

In this method, all loss functions are simply optimized by employing the Adam and L-BFGS algorithm, in which Adam optimization algorithm is a variant of the traditional stochastic gradient descent, whereas L-BFGS optimization algorithm is a full-batch gradient descent optimization algorithm based on a quasi-Newton method. [62,63]. Especially, the scalable parameters in the adaptive activation function are initialized generally as $n = 10$, $a_i^d = 0.1$, unless otherwise specified. In addition, we select relatively simple multi-layer perceptrons (i.e., feedforward NNs) with the Xavier initialization and the hyperbolic tangent (tanh) as activation function. All the codes in this article is based on Python 3.7 and Tensorflow 1.15, and all numerical experiments reported here are run on a DELL Precision 7920 Tower computer with 2.10 GHz 8-core Xeon Silver 4110 processor, 64 GB memory and 11 GB NVIDIA GeForce GTX 1080 Ti video card.

3. Data-driven vector solitons and breathers of the Manakov system

In this section, we will focus on the data-driven vector solitons and breather for the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ by means of 9 hidden layers deep IPINN with 40 neurons per layer, thus the physics-informed parts Eq. (4) of the IPINN for Manakov system (1) become the following formula

$$\begin{aligned} f_u &= -v_t + u_{xx} + 2u(u^2 + v^2 + m^2 + n^2), \\ f_v &= u_t + v_{xx} + 2v(u^2 + v^2 + m^2 + n^2), \\ f_m &= -n_t + m_{xx} + 2m(u^2 + v^2 + m^2 + n^2), \\ f_n &= m_t + n_{xx} + 2n(u^2 + v^2 + m^2 + n^2). \end{aligned} \tag{11}$$

The vector N -soliton and breather of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ have been derived by the Riemann-Hilbert method [55], the general vector form of N -soliton solution can be expressed as

$$\begin{bmatrix} q_1(x, t) \\ q_2(x, t) \end{bmatrix} = 2i \sum_{j,k=1}^N \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} e^{\theta_j - \theta_k} (M^{-1})_{jk}, \tag{12}$$

where $\theta_l = -i\zeta_l x - 2i\zeta_l^2 t$, and α_l, β_l and ζ_l are arbitrary constants in complex space ($l = 1, 2, \dots, N$). The symbol “*” represents complex conjugate, and M is $N \times N$ matrix whose elements are given by

$$M_{jk} = \frac{1}{\zeta_j^* - \zeta_k} \left[e^{-(\theta_j^* + \theta_k)} + (\alpha_j^* \alpha_k + \beta_j^* \beta_k) e^{\theta_j^* + \theta_k} \right]. \tag{13}$$

Next, we will exhibit the vector one-soliton, two-soliton and breather for the Manakov system in more detail.

• Vector one-solitons

In order to get the vector one-soliton solution, we set $N = 1$ in the above formulae (12), and take $\alpha_1 = 1/4$, $\beta_1 = 1/4$ and $\zeta_1 = 1 + 2i$, then the vector one-soliton solution in the Manakov system can be rewritten as

$$\begin{aligned} q_{1,os} &= \frac{16e^{2i(6t-x)}}{8e^{-4x-16t} + e^{4x+16t}}, \\ q_{2,os} &= \frac{16e^{2i(6t-x)}}{8e^{-4x-16t} + e^{4x+16t}}. \end{aligned} \tag{14}$$

• Vector two-solitons

When $N = 2$ and $\text{Re} \zeta_1 \neq \text{Re} \zeta_2$, solution (12) describe the collision of two one-soliton, to illustrate, we take $\zeta_1 = 0.1 + 0.7i$, $\zeta_2 = -0.1 + 0.4i$, $\alpha_1 = \alpha_2 = 1$, $\beta_1 = 1/4$, $\beta_2 = 0$ and the corresponding solution are derived as

$$\begin{aligned} q_{1,ts} &= \frac{\Omega}{3584 \cos\left(\frac{33}{25}t - \frac{2}{5}x\right) - 2000e^{-\frac{28}{25}t - \frac{8}{5}x} - 333e^{\frac{28}{25}t + \frac{14}{5}x} - 208e^{-\frac{28}{25}t - \frac{14}{5}x} - 2125e^{\frac{28}{25}t + \frac{14}{5}x}}, \\ q_{2,ts} &= \frac{\Delta}{3584 \cos\left(\frac{33}{25}t - \frac{2}{5}x\right) - 2000e^{-\frac{28}{25}t - \frac{8}{5}x} - 333e^{\frac{28}{25}t + \frac{14}{5}x} - 208e^{-\frac{28}{25}t - \frac{14}{5}x} - 2125e^{\frac{28}{25}t + \frac{14}{5}x}}, \end{aligned} \tag{15}$$

with

$$\begin{aligned} \Omega &= \frac{8}{5}i \left(1036ie^{\frac{48}{25}it - \frac{1}{5}ix - \frac{8}{25}t + \frac{4}{5}x} + 1036ie^{\frac{48}{25}it - \frac{1}{5}ix + \frac{8}{25}t - \frac{4}{5}x} \right. \\ &\quad - 464ie^{\frac{2}{5}it + \frac{1}{5}ix - \frac{14}{25}t - \frac{7}{5}x} - 339ie^{\frac{2}{5}it + \frac{1}{5}ix + \frac{14}{25}t + \frac{7}{5}x} \\ &\quad - 448e^{\frac{48}{25}it - \frac{1}{5}ix + \frac{8}{25}t - \frac{4}{5}x} + 448e^{\frac{48}{25}it - \frac{1}{5}ix - \frac{8}{25}t + \frac{4}{5}x} \\ &\quad \left. - 448e^{\frac{2}{5}it + \frac{1}{5}ix + \frac{14}{25}t + \frac{7}{5}x} \right), \\ \Delta &= \frac{56}{5}i \left(125ie^{\frac{48}{25}it - \frac{1}{5}ix - \frac{8}{25}t + \frac{4}{5}x} + 37ie^{\frac{48}{25}it - \frac{1}{5}ix + \frac{8}{25}t - \frac{4}{5}x} \right. \\ &\quad \left. - 88ie^{\frac{2}{5}it + \frac{1}{5}ix + \frac{14}{25}t + \frac{7}{5}x} - 16e^{\frac{48}{25}it - \frac{1}{5}ix + \frac{8}{25}t - \frac{4}{5}x} - 16e^{\frac{2}{5}it + \frac{1}{5}ix + \frac{14}{25}t + \frac{7}{5}x} \right). \end{aligned}$$

• Vector breathers

Specially, if $N = 2$ and $\text{Re} \zeta_1 = \text{Re} \zeta_2$, one can obtain the vector breather solution for the Manakov system, by substituting $\zeta_1 = 0.7i$, $\zeta_2 = 0.4i$, $\alpha_1 = \alpha_2 = 1$, $\beta_1 = 1/4$, $\beta_2 = 0$ into Eq. (12), we have

$$\begin{aligned} q_{1,bs} &= \frac{88}{5} \frac{84e^{\frac{48}{25}it + \frac{4}{5}x} + 84e^{\frac{48}{25}it - \frac{4}{5}x} - 48e^{-\frac{7}{5}x + \frac{16}{25}it} - 37e^{\frac{7}{5}x + \frac{16}{25}it}}{-3584 \cos\left(\frac{33}{25}t\right) + 1936e^{-\frac{2}{5}x} + 265e^{\frac{1}{5}x} + 144e^{-\frac{1}{5}x} + 2057e^{\frac{2}{5}x}}, \\ q_{2,bs} &= \frac{616}{5} \frac{11e^{\frac{48}{25}it + \frac{4}{5}x} + 3e^{\frac{48}{25}it - \frac{4}{5}x} - 8e^{\frac{7}{5}x + \frac{16}{25}it}}{-3584 \cos\left(\frac{33}{25}t\right) + 1936e^{-\frac{2}{5}x} + 265e^{\frac{1}{5}x} + 144e^{-\frac{1}{5}x} + 2057e^{\frac{2}{5}x}}. \end{aligned} \tag{16}$$

Furthermore, we can also obtain higher-order vector solitons and breathers by taking different N from Eq. (12). Accordingly, the form and dynamic behavior of higher-order vector solitons and breathers are more complex, with more parameters and stronger adjustability. In this paper, we only consider data-driven vector one-soliton, two-soliton and breather, and the higher-order case is similar.

3.1. Data-driven vector one-solitons

In this subsection, we will focus on the data-driven vector one-solitons with the initial conditions $q_r^0(x)$ and Dirichlet boundary conditions $q_r^{lb}(t)$ and $q_r^{ub}(t)$, $r = 1, 2$. Here one can take $[L_0, L_1]$ and $[T_0, T_1]$ in Eq. (1) as $[-3.0, 3.0]$ and $[-0.2, 0.2]$, respectively. Then, we obtain the corresponding initial conditions as shown below

$$q_r^0(x) = q_{r,os}(x, -0.2), x \in [-3.0, 3.0], \tag{17}$$

and the Dirichlet boundary conditions

$$q_r^{lb}(t) = q_{r,os}(-3.0, t), q_r^{ub}(t) = q_{r,os}(3.0, t), t \in [-0.2, 0.2], r = 1, 2. \tag{18}$$

We employ the traditional finite difference scheme on even grids in Matlab to simulate vector one-solitons which contains the initial datas

(17) and boundary datas (18) to acquire the original training data. More specifically, we divide spatial region $[-3.0,3.0]$ into 2500 points and temporal region $[-0.2,0.2]$ into 1000 points, the vector one-solitons (14) is discretized into 1000 snapshots accordingly. We generate a smaller training dataset containing initial-boundary data by randomly extracting $N_q = 2500$ from original dataset and $N_f = 40000$ collocation points which are generated by the Latin Hypercube Sampling method (LHS) [64]. After giving a dataset of initial and boundary points, the latent vector one-solitons $q_r(x,t)$ ($r = 1,2$) have been successfully learned by tuning all learnable parameters of the 9 layers IPINN with 40 neurons per layer and regulating the loss function, in which we utilize the 50,000 steps Adam firstly and subsequently use 15,134 steps L-BFGS optimizations for minimizing the loss function (7). The relative \mathbb{L}_2 errors of the IPINN model are $3.813803e^{-2}$ for $q_1(x,t)$ and $3.754237e^{-2}$ for $q_2(x,t)$, the total number of iterations is 65,134.

Fig. 2–4 display the deep learning results of vector one-solitons based on the IPINN related to the initial boundary value problem (17) and (18) of the Manakov system (1). Specifically, the density plots with the corresponding peak scale for diverse dynamics which contain exact dynamics, learned dynamics and error dynamics have been exhibited in detail, and the sectional drawings of one soliton solutions $q_r(x,t)$ ($r = 1,2$) at different moments arising from the IPINN are displayed in Fig. 2. From the bottom panels of Fig. 2(a) and (b), one can indicate vector one-solitons propagate from right to left along the x -axis, and show that the amplitudes of solitons are constant with the development of time t . Fig. 3 exhibit the three-dimensional plots and corresponding contour maps of the predicted vector one-soliton solutions $q_r(x,t)$ stemmed from the IPINN. The loss function curve figures account for the vector one-solitons $q_r(x,t)$ have been given out in Fig. 4, the left panel (a) of Fig. 4 demonstrates that Adam is an optimizer with oscillatory loss function curves, the left panel (b) in Fig. 4 illustrates L-BFGS is an optimization algorithm with linear loss function curves. Due to the $Loss_a$ possesses unique topology structure (10), the $Loss_a$ curves decrease slowly around 0.001 and have a strong stability in

Fig. 4. From Fig. 4 (a), we will find that the loss function curves of $Loss_{q_1}$, $Loss_{q_2}$, $Loss_{f_1}$ and $Loss_{f_2}$ overlap, which is caused by the consistency of the initial-boundary value conditions of $q_1(x,t)$ and $q_2(x,t)$. By the way, according to Eqs. (12)–(14), one can observe that we can take unequal values of α_1 and β_1 in Eq. (12) to adjust the amplitude and initial boundary value conditions for the vector one-solitons. Of course, for simplicity, we take the case of equal amplitude in this subsection.

3.2. Data-driven vector two-solitons

In what follows, we will consider the initial-boundary value problem of the Manakov system for obtaining the data-driven vector two-solitons by applying the multilayer IPINN. Similarly, taking $[L_0, L_1]$ and $[T_0, T_1]$ in Eq. (1) as $[-6.0,6.0]$ and $[-3.0,3.0]$ respectively, we derive the corresponding initial condition $q_r^0(x)$ and Dirichlet boundary conditions $q_r^{lb}(t)$ and $q_r^{ub}(t)$, $r = 1, 2$ as shown in the following formulas

$$q_r^0(x) = q_{r,ts}(x, -3.0), x \in [-6.0, 6.0], \tag{19}$$

and

$$q_r^{lb}(t) = q_{r,ts}(-6.0, t), q_r^{ub}(t) = q_{r,ts}(6.0, t), t \in [-3.0, 3.0], r = 1, 2. \tag{20}$$

By means of Matlab, we discretize the Eq. (15) by utilizing the traditional finite difference scheme on even grids, and obtain the original training data which only contains initial data (19) and boundary data (20) by dividing the spatial region $[-6.0,6.0]$ into 2000 points and the temporal region $[-3.0,3.0]$ into 1500 points, the remaining data will be used to obtain training errors by comparing with predicted vector two-solitons. After that, we generate a smaller training dataset containing initial-boundary data by randomly extracting $N_q = 2000$ from original training dataset and $N_f = 30000$ collocation points produced via LHS in the corresponding spatiotemporal region. Then, the latent

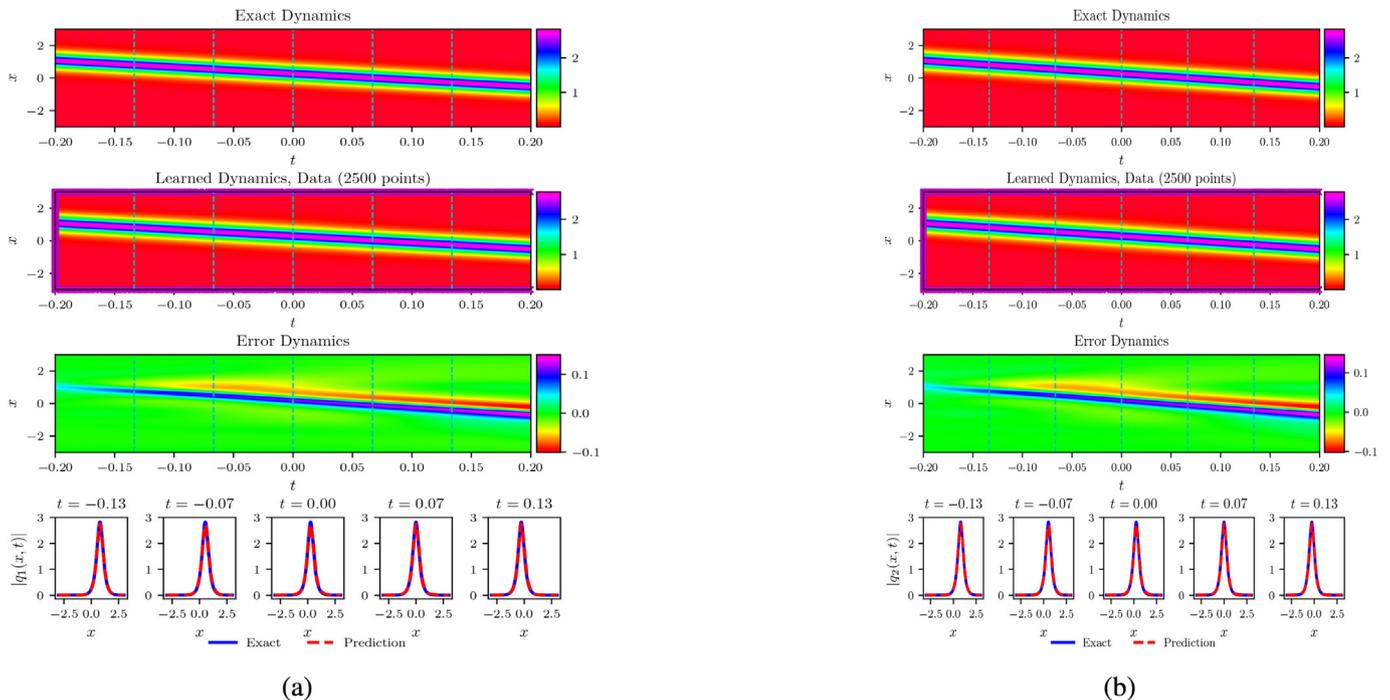


Fig. 2. The vector one-solitons $q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2500$ which have been shown by using mediumchid “ \times ” in learned dynamics, and $N_f = 40000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector one-solitons $q_r(x,t)$ with five distinct training moments $t = -0.13, -0.07, 0.00, 0.07$ and 0.13 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector one-solitons $q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the one-soliton $q_1(x,t)$; (b) The density plots and sectional drawings for the one-soliton $q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

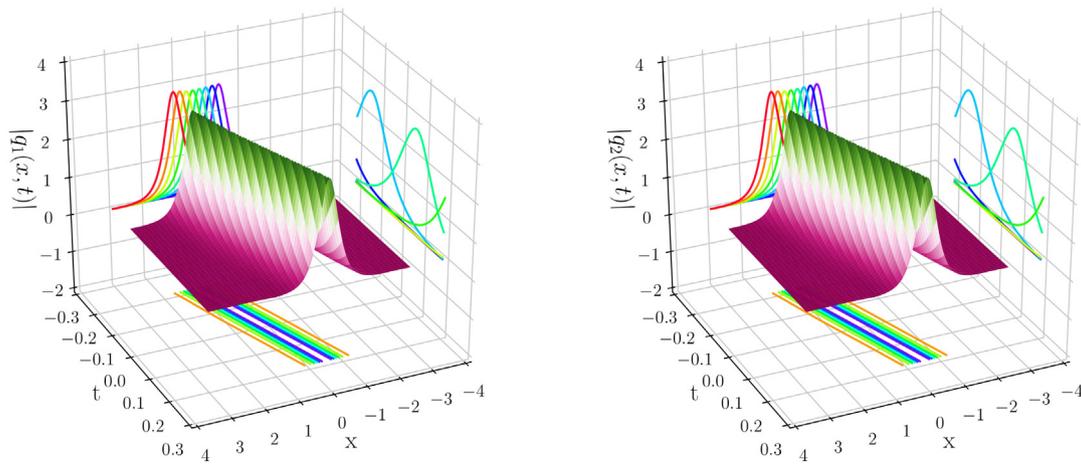


Fig. 3. The three-dimensional plots with contour map on three planes of the predicted vector one-solitons $q_r(x,t)$ ($r = 1,2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x,t)$; (Right side panel) The 3D plot for the $q_2(x,t)$.

vector two-solitons $q_r(x,t)$ ($r = 1,2$) have been successfully learned by imposing a 9-hidden-layer IPINN with 40 neurons per layer, and the related loss functions are optimized through 20,000 Adam iterations and 26,239 L-BFGS iterations. The relative L_2 errors of the IPINN model are $2.678111e^{-2}$ for $q_1(x,t)$ and $2.356420e^{-2}$ for $q_2(x,t)$, the total number of iterations is 46,239.

Figs. 5–7 display the training results of the vector two-solitons $q_r(x,t)$ ($r = 1,2$) based on the IPINN related to the initial-boundary value problem (19) and (20) of the Manakov system (1). Fig. 5 depicts various dynamic density plots and sectional drawing at different moments, in which the panel (a) corresponds to two-soliton solution $q_1(x,t)$ and panel (b) corresponds to two-soliton solution $q_2(x,t)$ for the Manakov system (1). As we can see from the bottom panels in Fig. 5 (a) and (b), the $q_1(x,t)$ soliton before collision has two components, but the $q_2(x,t)$ soliton before collision has only a component. However, after collision its another component of the $q_2(x,t)$ soliton appears when this soliton has re-emerged on the right side by combining learned dynamics density plot in Fig. 5 and 3D plots in Fig. 6, it also means that the two-soliton solution $q_2(x,t)$ satisfies a soliton fission process. Of course, the total energy of each component soliton is conserved before and after collision. Fig. 6 exhibits the three-dimensional plots with projected contours on three planes for the

learning vector two-solitons $q_r(x,t)$, which corresponds to the two density plots of learned dynamics in Fig. 5. Fig. 7 showcases curve plots of the loss function after 20,000 Adam optimization iterations and 26,239 L-BFGS optimization iterations in IPINN framework. Different from the vector one-soliton solutions in Section 3.1, the loss function curves of Adam optimization for the vector two-soliton solutions are more hierarchical, which is caused by the different initial-boundary value conditions of $q_1(x,t)$ and $q_2(x,t)$. Obviously, from the bottom panels (a) and (b) of Fig. 5, one can find that the amplitude of $q_1(x,t)$ is always greater than that of $q_2(x,t)$ at a fixed moment, this also causes the loss function values $Loss_{q_1}$ and $Loss_{f_1}$ to be greater than $Loss_{q_2}$ and $Loss_{f_2}$ respectively during the Adam optimization iteration. On the other hand, the loss function curves $Loss_{q_1}$ and $Loss_{q_2}$ of L-BFGS optimization are missing after a certain number of iterations, that is because the loss function value in this part is less than $1e^{-6}$, which is beyond the statistical range of “% f” in Python 3.7.

3.3. Data-driven vector breathers

It is well known that the breather solution is a special two-soliton solution, so it is also called the soliton bound state solution. Next, we will consider the data-driven vector breather solutions for the Manakov

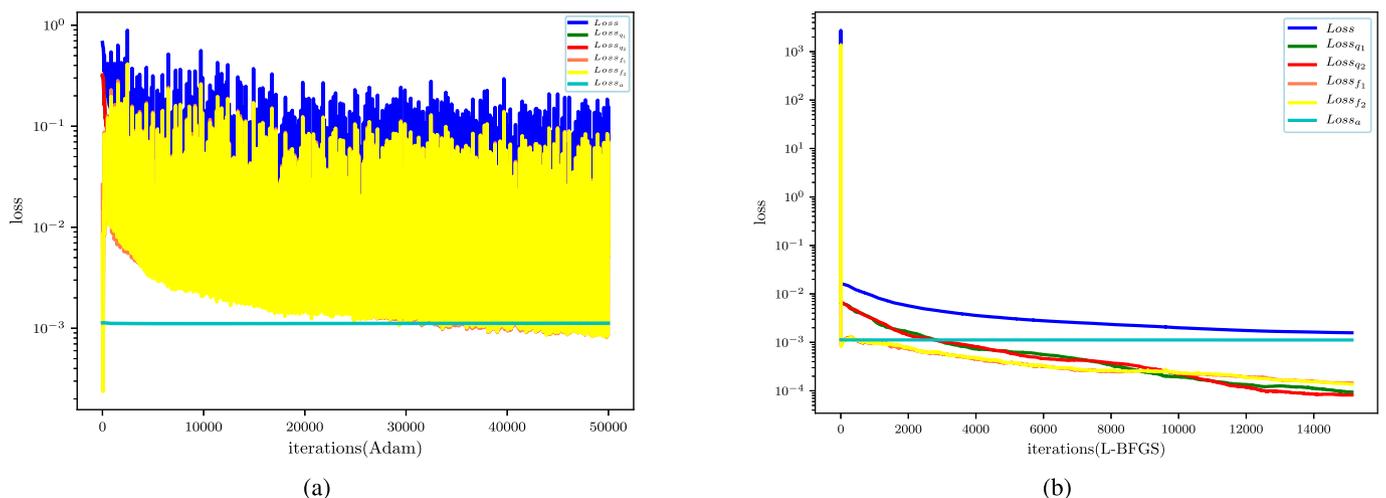


Fig. 4. The loss function curve figures of the vector one-solitons $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 50,000 steps Adam and 15,134 steps L-BFGS optimizations: (a) The loss function curve for the 50,000 Adam optimization iterations; (b) The loss function curve for the 15,134 L-BFGS optimization iterations.

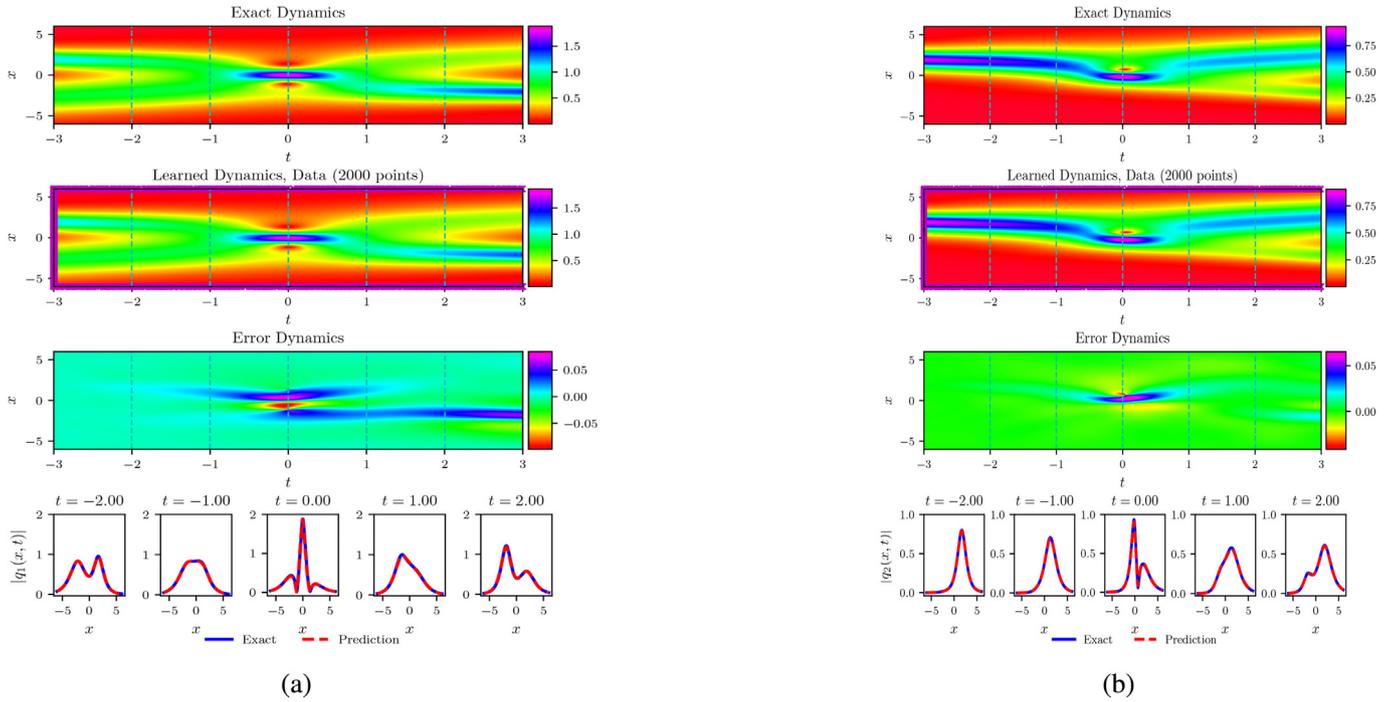


Fig. 5. The vector two-solitons $q_r(x, t)$ ($r = 1, 2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “ \times ” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector two-solitons $q_r(x, t)$ with five distinct training moments $t = -2.00, -1.00, 0.00, 1.00$ and 2.00 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector two-solitons $q_r(x, t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the two-soliton $q_1(x, t)$; (b) The density plots and sectional drawings for the two-soliton $q_2(x, t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

system by utilizing the initial-boundary value conditions and the 9-layer IPINN with 40 neurons per layer. Similarly, considering the initial conditions $q_r^0(x)$ ($r = 1, 2$) and Dirichlet boundary conditions $q_r^b(t)$ and $q_r^{ub}(t)$ of Eq. (1) arising from the exact breathers Eq. (16), we take $[L_0, L_1]$ and $[T_0, T_1]$ in Eq. (1) as $[-6.0, 6.0]$ and $[-3.0, 3.0]$, respectively. Therefore, the corresponding initial conditions can be written as bellows

$$q_r^0(x) = q_{r,bs}(x, -3.0), x \in [-6.0, 6.0], \tag{21}$$

and the Dirichlet boundary conditions become

$$q_r^b(t) = q_{r,bs}(-6.0, t), q_r^{ub}(t) = q_{r,bs}(6.0, t), t \in [-3.0, 3.0], r = 1, 2. \tag{22}$$

Similarly, discretizing Eq. (16) with the aid of the traditional finite difference scheme on even grids in Matlab, and we obtain the original training data which contain initial data (21) and boundary data (22) by dividing the spatial region $[-6.0, 6.0]$ into 2000 points and the temporal region $[-3.0, 3.0]$ into 1000 points. Then, one can generate a smaller training dataset that containing initial-boundary data by randomly extracting $N_q = 2000$ from original dataset and $N_f = 30000$ collocation points which are produced by the LHS. After that, the latent vector breathers $q_r(x, t)$ have been successfully learned by tuning all learnable parameters of the IPINN and regulating the loss function. The model of IPINN achieved relative \mathbb{L}_2 error of $1.631581e^{-2}$ for the breather solution $q_1(x, t)$ and relative \mathbb{L}_2 error of $1.442727e^{-2}$ for the breather solution $q_2(x, t)$, and the total number of

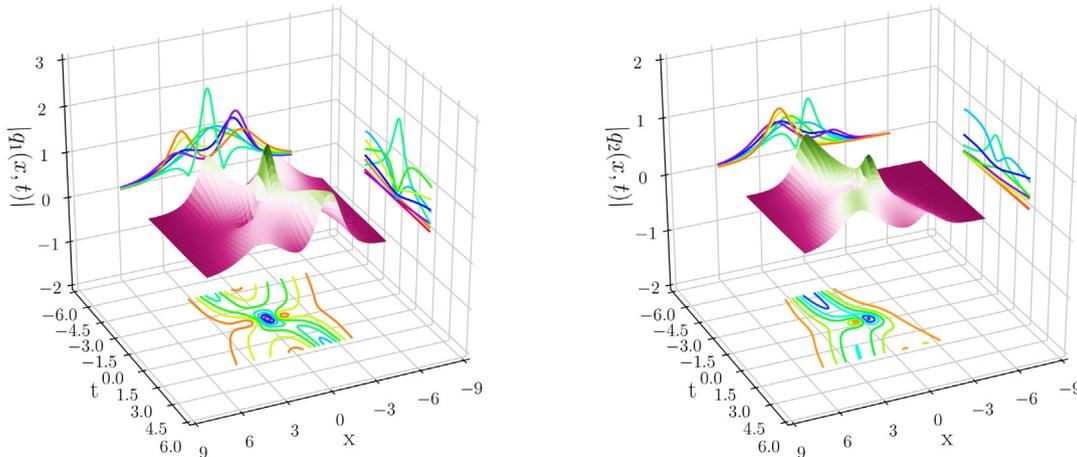


Fig. 6. The three-dimensional plots with contour map on three planes of the predicted vector two-solitons $q_r(x, t)$ ($r = 1, 2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x, t)$; (Right side panel) The 3D plot for the $q_2(x, t)$.

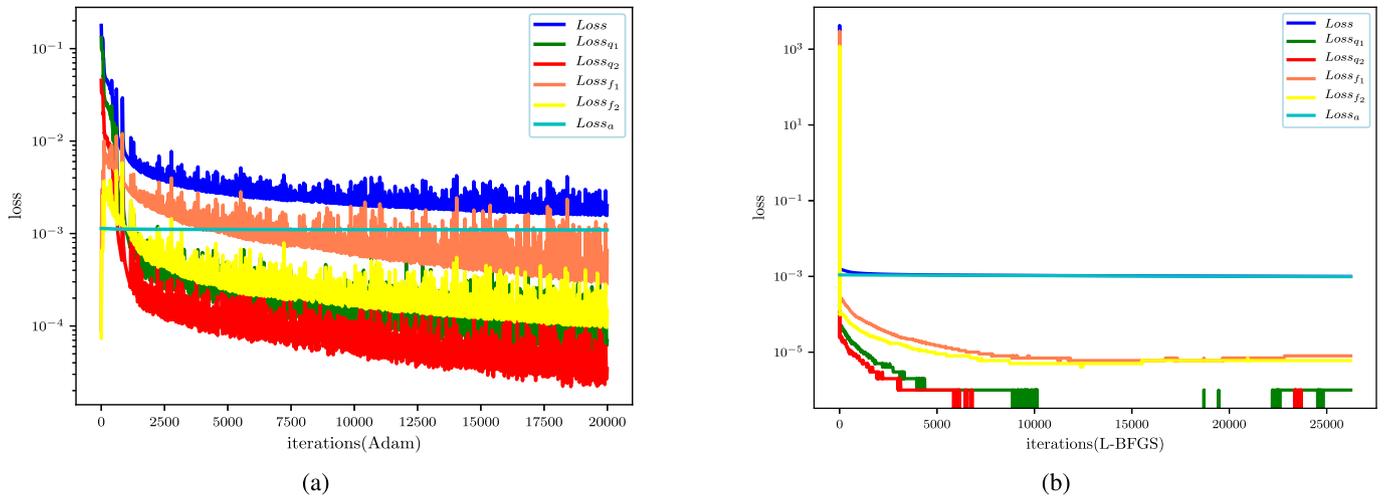


Fig. 7. The loss function curve figures of the vector two-solitons $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 26,239 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 26,239 L-BFGS optimization iterations.

iterations which contain 20,000 Adam optimizations and 22,374 L-BFGS optimizations is 42,374.

Figs. 8–10 display the deep learning results of the vector breathers $q_r(x,t)$ ($r = 1,2$) based on the IPINN related to the initial-boundary value problem (21) and (22) of the Manakov system (1). The dynamics behaviors which contain exact, learned, error dynamics of the vector breathers are shown in Fig. 8. Compared with the dynamics behaviors of vector two-solitons in the bottom panels of Fig. 5, one can notice that the shapes of the vector breathers are the same at $t = -1$ ($t = -2$) and $t = 1$ ($t = 2$) in the bottom panels of Fig. 8, which are different from the vector two solitons. From the learned dynamics density plots

in Fig. 8, it can be seen that this is a localized periodic oscillation solution, whose shape is similar to respiratory fluctuation, so it is also called “breather”. Fig. 9 exhibits three-dimensional plots of the predicted vector breathers $q_r(x,t)$ for the Eq. (1) result from the IPINN. The graph in Fig. 10(a) is similar to that of the case of vector two-solitons in Fig. 7(a), and there are more missing parts of the $Loss_{q1}$ and $Loss_{q2}$ in Fig. 10(b), which also indicate that the optimization for the vector breathers is better than the vector two-solitons arising from the L-BFGS optimization iterations.

In this section, based on the initial-boundary value problem of the Manakov system, we discuss the data-driven vector one-solitons, two-

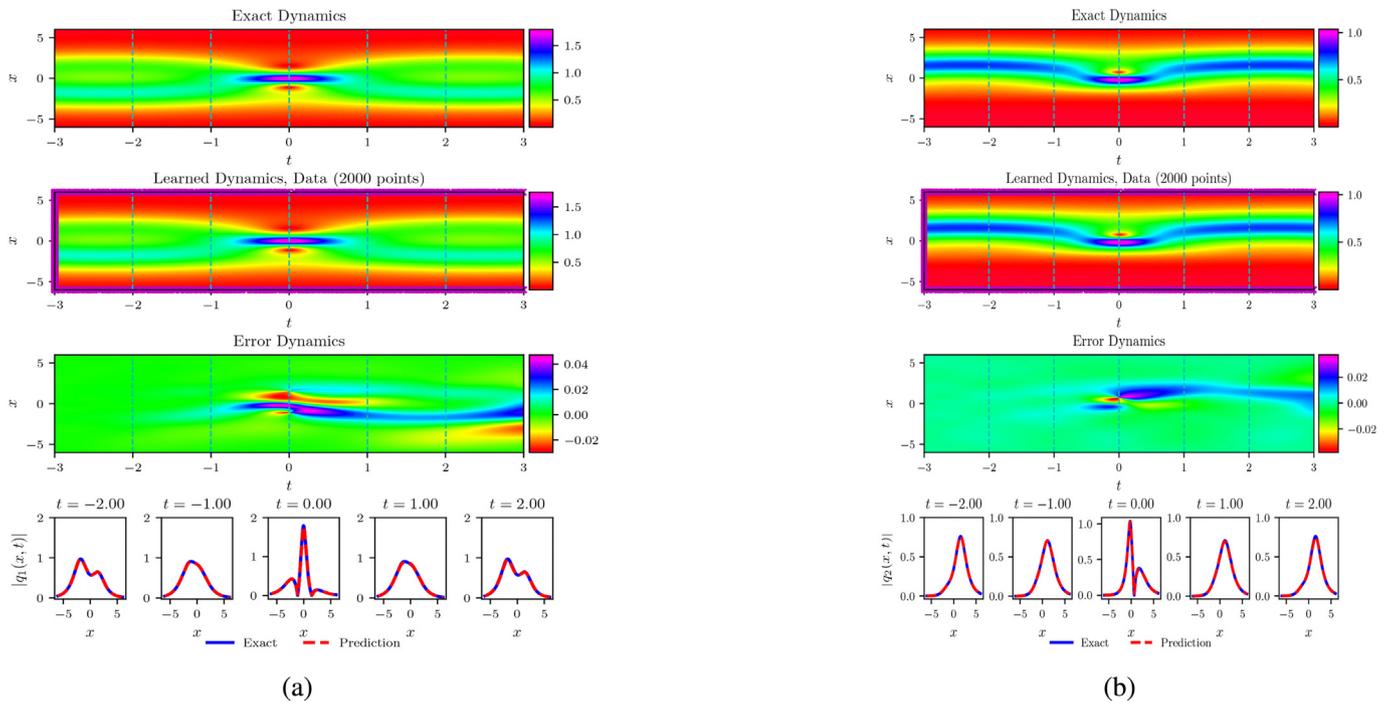


Fig. 8. The vector breathers $q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “ \times ” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector breathers $q_r(x,t)$ with five distinct training moments $t = -2.00, -1.00, 0.00, 1.00$ and 2.00 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector breathers $q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the breather $q_1(x,t)$; (b) The density plots and sectional drawings for the breather $q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

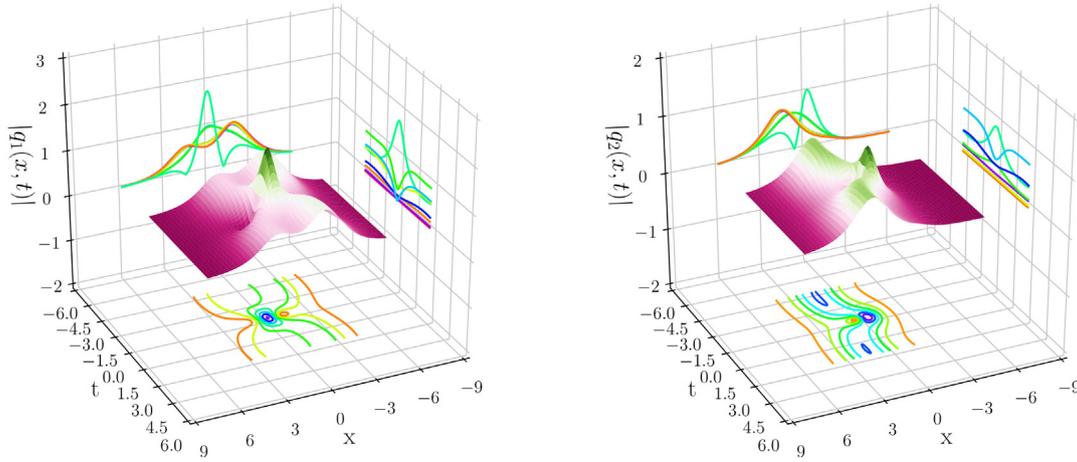


Fig. 9. The three-dimensional plots with contour map on three planes of the predicted breather solutions $q_r(x,t)$ ($r = 1,2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x,t)$; (Right side panel) The 3D plot for the $q_2(x,t)$.

solitons and breathers via the IPINN model. However, for more data-driven vector N -solitons and N -breathers, such as the vector two-breathers which derived from the vector four-solitons, we can also recover them through the IPINN framework. Of course, it will require us to take more initial-boundary value data points and residual configuration points, establish deeper and wider NNs, and demand stronger CPU and GPU of computer. Table 1 exhibits relative \mathbb{L}_2 norm errors of three different types of localized waves by means of IPINN.

4. Data-driven vector rogue waves of the Manakov system

In various complex systems, the research of vector RWs has more important physical significance, and the multi-component model has more appropriate practical significance. In this section, we will focus on the data-driven vector RWs of the Manakov system (1) with $\lambda_1 = 1$ and $\lambda_2 = 2$ based on the IPINN. That is the physics-informed parts are consistent with Eq. (11) in Section 3. The abundant interaction solutions between RWs and various other types of waves and N -order RWs with parameter adjustment have been obtained by means of Darboux transformation. Next, in order to compare the errors of vector RWs learned by using the IPINN, we give out the exact solution expressions of several common vector RWs.

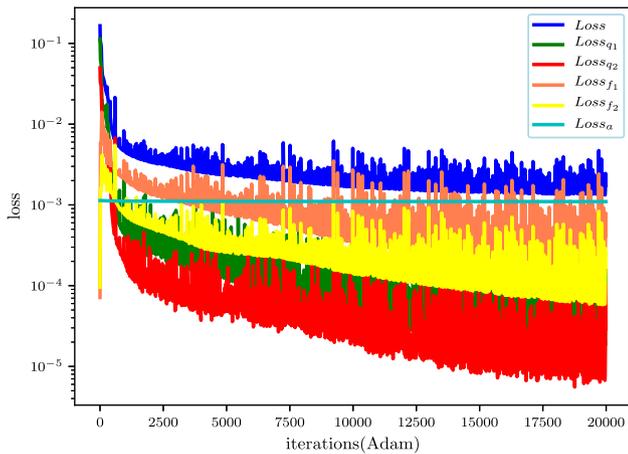
• Vector Rogue Waves I

One can obtain the explicit vector RWs of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ as following

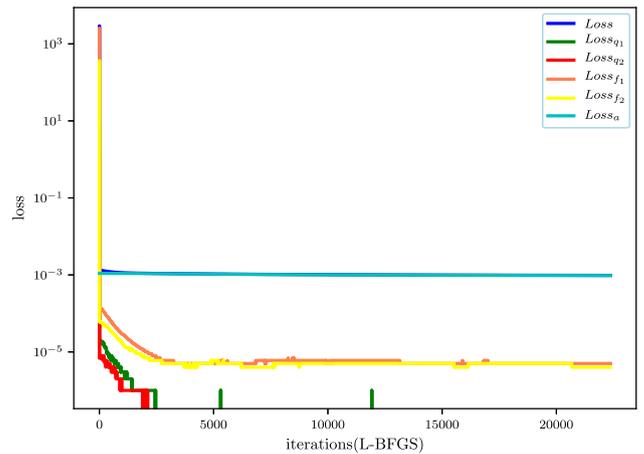
$$\begin{aligned}
 q_{1,rw1}(x,t) &= -\frac{3}{10} \cdot \frac{\Omega_1}{900t\sqrt{3} + 1500\sqrt{3}x - 972t^2 - 810tx - 675x^2 - 3125}, \\
 q_{2,rw1}(x,t) &= \frac{3}{10} \cdot \frac{\Omega_2}{900t\sqrt{3} + 1500\sqrt{3}x - 972t^2 - 810tx - 675x^2 - 3125}, \\
 \Omega_1 &= (972i\sqrt{3}t^2 + 810i\sqrt{3}tx + 675i\sqrt{3}x^2 - 4050it - 3375ix + 450t\sqrt{3} - 2625\sqrt{3}x + 972t^2 + 810tx + 675x^2 + 5000)e^{\frac{36i}{5}it}, \\
 \Omega_2 &= (972i\sqrt{3}t^2 + 810i\sqrt{3}tx + 675i\sqrt{3}x^2 - 3375ix + 3600t\sqrt{3} + 2625\sqrt{3}x - 972t^2 - 810tx - 675x^2 - 5000)e^{\frac{36i}{5}(i - 5x+9t)},
 \end{aligned}
 \tag{23}$$

the vector RWs of the Manakov system have been derived by Guo and Ling [57], this kind of RW solutions are similar to the first-order RW solutions for the NLS [8].

• Vector Rogue Waves II



(a)



(b)

Fig. 10. The loss function curve figures of the vector breathers $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 22,374 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 22,374 L-BFGS optimization iterations.

Table 1
Relative \mathbb{L}_2 errors of three different localized waves types in IPINN model.

Component	Wave types		
	Vector one-solitons	Vector two-solitons	Vector breathers
Component q_1	3.813803e-02	2.678111e-02	1.631581e-02
Component q_2	3.754237e-02	2.356420e-02	1.442727e-02

Furthermore, in addition to the above vector RWs I, Guo and Ling also obtained a more complex vector RWs II [57], the form of the explicit vector RWs II is as follows

$$\begin{aligned}
 q_{1,rw2}(x, t) &= -\frac{3}{10} \frac{\Delta_1}{\Xi}, \\
 q_{2,rw2}(x, t) &= \frac{3}{10} \frac{\Delta_2}{\Xi},
 \end{aligned} \tag{24}$$

where

$$\begin{aligned}
 \Delta_1 &= (-3125000 + 524880i\sqrt{3}tx + 656100i\sqrt{3}t^2x^2 + 364500i\sqrt{3}tx^3 + 5062500it\sqrt{3}x + \\
 &6682500t^2 - 2187000\sqrt{3}t^2x - 1215000\sqrt{3}tx^2 + 562500t\sqrt{3} + 1012500tx + 5906250x^2 \\
 &- 1875000\sqrt{3}x - 15187500it^3 - 3936600it^3 - 1518750it^3 + 524880t^3x + 656100t^3x^2 + \\
 &364500t^3x^3 - 1181250\sqrt{3}x^3 - 145800\sqrt{3}t^3 + 314928i\sqrt{3}t^4 + 151875i\sqrt{3}x^4 + 4252500i\sqrt{3}t^2 \\
 &+ 843750i\sqrt{3}x^2 - 6561000it^2x - 3645000itx^2 + 314928t^4 + 151875x^4)e^{2it}, \\
 \Delta_2 &= (3125000 + 524880i\sqrt{3}tx + 656100i\sqrt{3}t^2x^2 + 364500i\sqrt{3}tx^3 - 1312200it^3 - 13972500t^2 \\
 &+ 4374000\sqrt{3}t^2x + 3037500\sqrt{3}tx^2 + 2812500t\sqrt{3} - 13162500tx - 5906250x^2 + 1875000\sqrt{3}x \\
 &- 3037500it\sqrt{3}x - 1518750it^3 - 524880t^3x - 656100t^3x^2 - 364500t^3x^3 + 1181250\sqrt{3}x^3 + \\
 &2770200\sqrt{3}t^3 + 15187500it + 314928i\sqrt{3}t^4 + 151875i\sqrt{3}x^4 + 843750i\sqrt{3}x^2 - 4374000it^2x - \\
 &1822500itx^2 - 314928t^4 - 151875x^4 - 607500i\sqrt{3}t^2)e^{2it(-5x+9t)}, \\
 \Xi &= 1020600\sqrt{3}t^3 + 2187000\sqrt{3}t^2x + 1215000\sqrt{3}tx^2 + 675000\sqrt{3}x^3 - 314928t^4 - \\
 &524880t^3x - 656100t^3x^2 - 364500t^3x^3 - 151875x^4 + 1125000t\sqrt{3} + 1875000\sqrt{3}x \\
 &- 6682500t^2 - 4050000tx - 3375000x^2 - 1562500.
 \end{aligned}$$

The vector RWs II are different behaviors to RWs of NLS [8], and different form Ref. [19] by a symmetry analysis, which is nothing but the RWs of NLS an different altitude.

• Vector Rogue Waves III

In order to obtain the interaction solutions between dark-bright solitons and RWs, authors also derived this interesting solutions by Darboux transformation method in Ref. [57], as shown below

$$\begin{aligned}
 q_{1,rw3}(x, t) &= \frac{\Lambda}{\Pi}, \\
 q_{2,rw3}(x, t) &= \frac{4(ix - 1 - i + 4t - x)e^{-\frac{2}{3}t + \frac{1}{3}x + 2it + ix}}{\Pi},
 \end{aligned} \tag{25}$$

where

$$\begin{aligned}
 \Lambda &= e^{it+ix}(8ie^{\frac{4}{3}t-\frac{2}{3}x}t - 16e^{\frac{4}{3}t-\frac{2}{3}x}t^2 + 8e^{\frac{4}{3}t-\frac{2}{3}x}tx - 2e^{\frac{4}{3}t-\frac{2}{3}x}x^2 - \\
 &4ie^{\frac{4}{3}t-\frac{2}{3}x} + 12e^{\frac{4}{3}t-\frac{2}{3}x}t - 2e^{\frac{4}{3}t-\frac{2}{3}x}x - e^{\frac{4}{3}t-\frac{2}{3}x} + 2e^{-\frac{4}{3}t+\frac{2}{3}x}), \\
 \Pi &= 16e^{\frac{4}{3}t-\frac{2}{3}x}t^2 - 8e^{\frac{4}{3}t-\frac{2}{3}x}tx + 2e^{\frac{4}{3}t-\frac{2}{3}x}x^2 - 12e^{\frac{4}{3}t-\frac{2}{3}x}t + 2e^{\frac{4}{3}t-\frac{2}{3}x}x + 3e^{\frac{4}{3}t-\frac{2}{3}x} + 2e^{-\frac{4}{3}t+\frac{2}{3}x}.
 \end{aligned}$$

This kind of solutions behaves like the nonlinear superposition for the RWs and dark-bright soliton. This vector RWs (25) reveal the mechanism and dynamic behavior of RWs produced in dark-bright soliton.

• Vector Rogue Waves IV

For the vector RWs IV, we will exhibit the dark-bright solitons together with single Peregrine solitons, which have been arrived at by means of the Darboux transformation theory [60], it can be expressed by the following formula

$$\begin{aligned}
 q_{1,rw4}(x, t) &= \frac{e^{2it}(1600it - 1600t^2 - 400x^2 + e^{-2x} + 300)}{1600t^2 + 400x^2 + e^{-2x} + 100}, \\
 q_{2,rw4}(x, t) &= \frac{(20 - 20i)(2x + 4it + 1)e^{-x+3it}}{1600t^2 + 400x^2 + e^{-2x} + 100}.
 \end{aligned} \tag{26}$$

Baronio obtained this solution by employing the Darboux dressing technique for the first time [58]. Later, Wang et al. also obtained this solution through the generalized Darboux transformation [60]. These results provide evidence of an attractive interaction between the dark-bright waves and RWs, and the observed behavior may also be interpreted as a mechanism of generation of one RW out of a slowly moving boomeronic soliton.

• Vector Rogue Waves V

Moreover, once the vector RWs IV are obtained, the vector RWs V are also derived correspondingly, as shown below

$$\begin{aligned}
 q_{1,rw5}(x, t) &= \frac{\Upsilon_1}{\sqrt{2}e^{-2\sqrt{2}x} + 3200t^2 + 400x^2 + 50}, \\
 q_{2,rw5}(x, t) &= \frac{\Upsilon_2}{\sqrt{2}e^{-2\sqrt{2}x} + 3200t^2 + 400x^2 + 50}.
 \end{aligned} \tag{27}$$

where

$$\begin{aligned}
 \Upsilon_1 &= 40i\sqrt{2}e^{-\sqrt{2}x+6it}x - 80 \cdot 2^{\frac{3}{2}}e^{-\sqrt{2}x+6it}t - 40\sqrt{2}e^{-\sqrt{2}x+6it}x - 80 \cdot 2^{\frac{3}{2}}ie^{-\sqrt{2}x+6it}t + 10 \cdot 2^{\frac{3}{2}} \\
 &e^{-\sqrt{2}x+6it} - 10 \cdot 2^{\frac{3}{2}}e^{-\sqrt{2}x+6it} + \sqrt{2}e^{4it-2\sqrt{2}x} + 1600ie^{4it}t - 3200e^{4it}t^2 - 400e^{4it}x^2 + 150e^{4it}, \\
 \Upsilon_2 &= -40i\sqrt{2}e^{-\sqrt{2}x+6it}x + 80 \cdot 2^{\frac{3}{2}}e^{-\sqrt{2}x+6it}t + 40\sqrt{2}e^{-\sqrt{2}x+6it}x + 80 \cdot 2^{\frac{3}{2}}ie^{-\sqrt{2}x+6it}t - 10 \cdot 2^{\frac{3}{2}} \\
 &e^{-\sqrt{2}x+6it} + 10 \cdot 2^{\frac{3}{2}}e^{-\sqrt{2}x+6it} + \sqrt{2}e^{4it-2\sqrt{2}x} + 1600ie^{4it}t - 3200e^{4it}t^2 - 400e^{4it}x^2 + 150e^{4it}.
 \end{aligned}$$

This is an interaction solutions obtained by the nonlinear superposition of breathers and RWs, which are more complex than the vector RWs IV [58,60]. Furthermore, due to the introduction of periodic oscillatory breather, the study of this solution plays an important role in the stability analysis of optical fiber transmission.

4.1. Data-driven vector rogue waves I

In this section, in order to recover the data-driven vector RWs I of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$, we will commit to introducing the initial boundary value conditions of the Manakov system to the 9-layer IPINN with 40 neurons per layer. Selecting $[L_0, L_1]$ and $[T_0, T_1]$ in Eq. (1) as $[-3.0, 6.0]$ and $[-2.0, 2.0]$ respectively, we can write the corresponding initial value conditions as follows

$$q_r^0(x) = q_{r,rw1}(x, -2.0), x \in [-3.0, 6.0], \tag{28}$$

and the Dirichlet boundary conditions become

$$q_r^b(t) = q_{r,rw1}(-3.0, t), q_r^{ub}(t) = q_{r,rw1}(6.0, t), t \in [-2.0, 2.0], r = 1, 2. \tag{29}$$

In order to obtain the original training data set of the above initial boundary value conditions (28) and (29), we discretize the exact vector RWs I (23) based on the finite difference method by dividing the spatial region $[-3.0, 6.0]$ into 2000 points and the temporal region $[-2.0, 2.0]$ into 1000 points in Matlab. Furthermore, in addition to the data set composed of the aforementioned initial boundary value conditions, the residual data set is used to calculate the \mathbb{L}_2 norm error by comparing with the predicted vector RWs I. After that, a smaller training dataset that containing initial-boundary data will be generated by randomly extracting $N_q = 2000$ from original dataset and $N_f = 30000$ collocation points which are produced by the LHS. According to 20,000 Adam iterations and 25,117 L-BFGS iterations, the latent vector

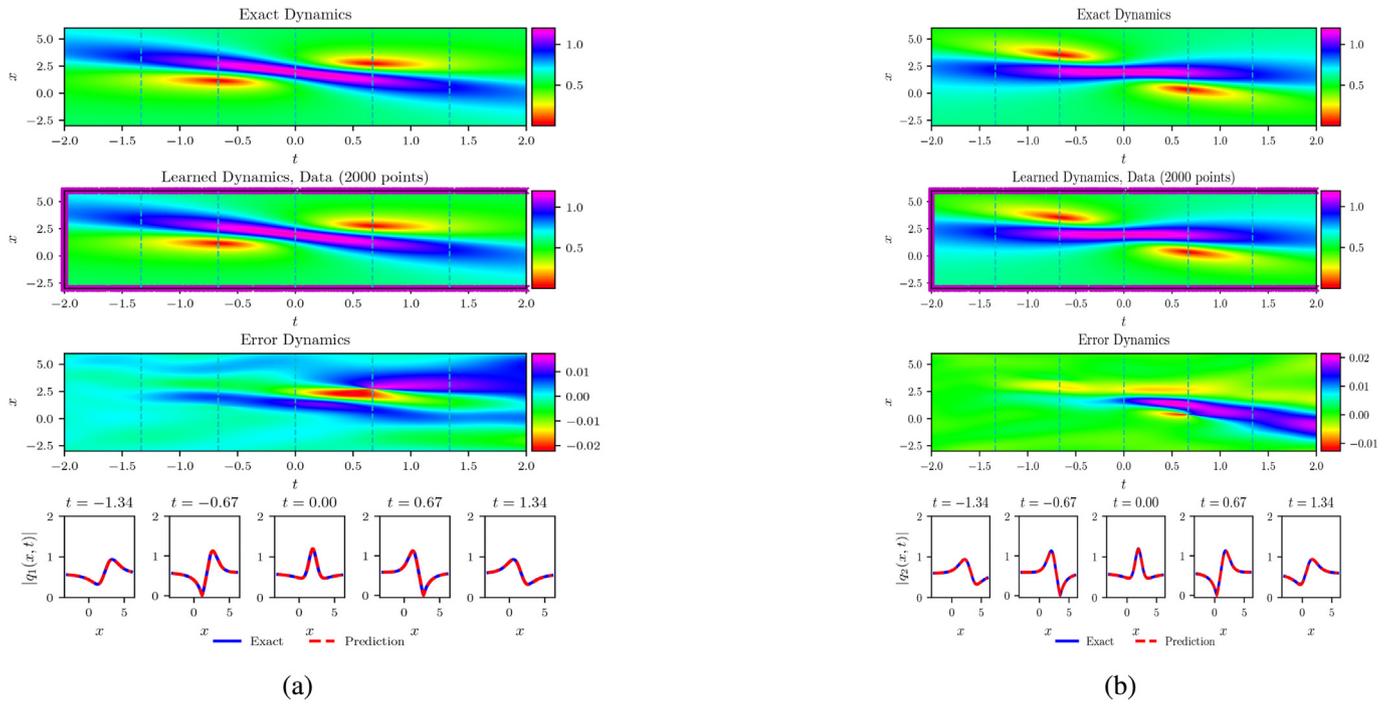


Fig. 11. The vector RWs $I q_r(x, t)$ ($r = 1, 2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “x” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector RWs $I q_1(x, t)$ with five distinct training moments $t = -1.34, -0.67, 0.00, 0.67$ and 1.34 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector RWs $I q_r(x, t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the RW $I q_1(x, t)$; (b) The density plots and sectional drawings for the RW $I q_2(x, t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

RWs $I q_r(x, t)$ have been successfully learned by employing the IPINN, and the network achieved relative \mathbb{L}_2 error of $7.505391e^{-3}$ for the RW $q_1(x, t)$ and relative \mathbb{L}_2 error of $7.988676e^{-3}$ for the RW $q_2(x, t)$, and the total number of iterations is 45,117.

Figs. 11–13 present the deep learning results of the vector RWs $I q_r(x, t)$ ($r = 1, 2$) based on the IPINN for the Manakov system with the initial-boundary value problem (28) and (29). Fig. 11 displays the exact, learned and error dynamics of the vector RWs I , and exhibits the sectional drawings which contain the learned and explicit RWs at five different moments. From the density plots of learned dynamics and profiles which reveal amplitude and error of exact and prediction RWs in Fig. 11, we can observe that the shapes are similar and amplitudes are identical for the vector RWs $q_1(x, t)$ and $q_2(x, t)$, but

the positions and orientations of the two-component RWs are different. The 3D plots of the predicted vector RWs $I q_r(x, t)$ have been shown through the IPINN in Fig. 12. Since the predicted vector RWs $I q_r(x, t)$ are similar in shape, the loss function curves of $Loss_{q_1}$ and $Loss_{q_2}$ ($Loss_{f_1}$ and $Loss_{f_2}$) are overlapped in Fig. 13(a). Due to the loss functions of $Loss_{q_1}$ and $Loss_{q_2}$ are small enough, some parts of the loss function curves are missing. However, when the number of iterations is 15,000 to 20,000 in Fig. 13(b), the loss function curve appears frequently from the missing state, which indicates that there are many oscillations in the loss function optimization at this stage by utilizing the L-BFGS optimizer, benefit from the introduction of loss function $Loss_a$, so that the total loss function $Loss$ decreases smoothly.

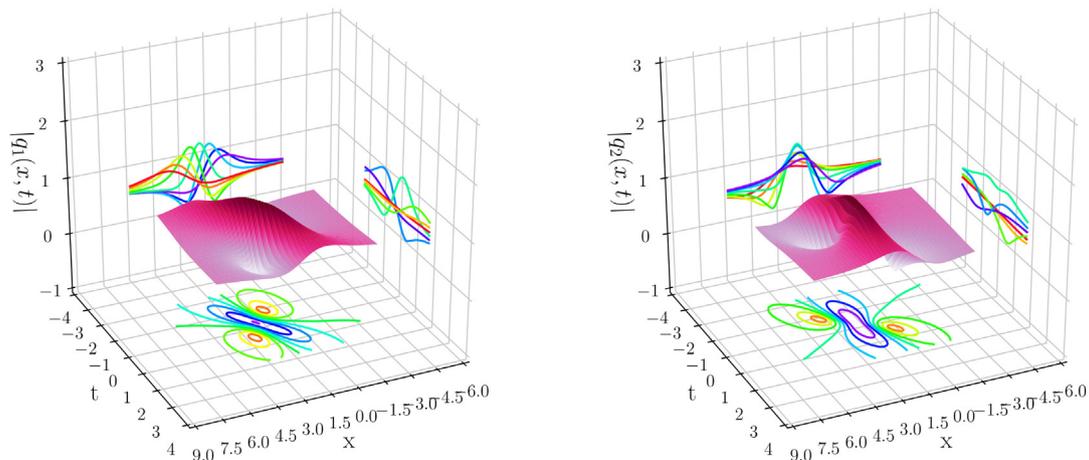


Fig. 12. The three-dimensional plots with contour map on three planes of the predicted vector RWs $I q_r(x, t)$ ($r = 1, 2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x, t)$; (Right side panel) The 3D plot for the $q_2(x, t)$.

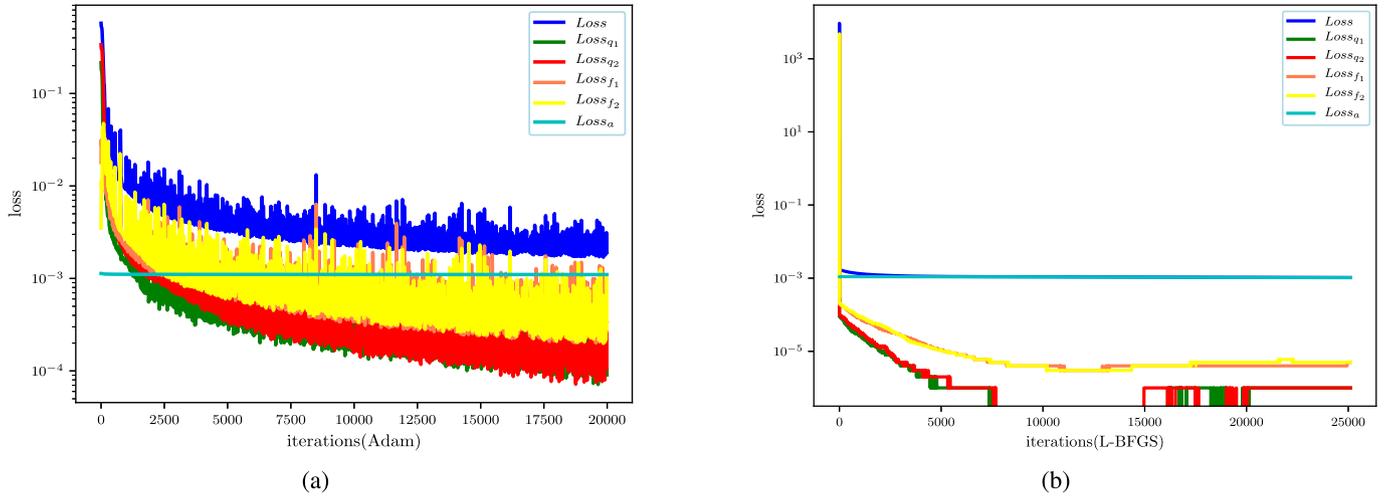


Fig. 13. The loss function curve figures of the vector RWs I $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 25,117 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 25,117 L-BFGS optimization iterations.

4.2. Data-driven vector rogue waves II

For recovering the data-driven vector RWs II of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ in the 9-layer IPINN with 40 neurons per layer, we consider the initial value conditions

$$q_r^0(x) = q_{r,rw2}(x, -1.0), x \in [-5.0, 10.0], \tag{30}$$

with the Dirichlet boundary conditions

$$q_r^l(t) = q_{r,rw2}(-5.0, t), q_r^{ub}(t) = q_{r,rw2}(10.0, t), t \in [-1.0, 1.0], r = 1,2. \tag{31}$$

With the aid of Matlab, the sampling points are selected by means of the finite difference method with spatial-temporal region $[-5.0, 10.0] \times [-1.0, 1.0]$, and the vector RWs (24) are discretized into $[2000 \times 1000]$ data points which contain initial-boundary value condition (30) and (31). Then, one can obtain the training dataset by randomly extracting $N_q = 2000$ from original initial-boundary value condition dataset and $N_f = 30000$ collocation points produced via the LHS in spatial-temporal region. Through 20,000 Adam iterations and 23,068 L-BFGS iterations to optimize loss function $\mathcal{L}(\bar{\Theta})$, the latent vector RWs II $q_r(x,t)$ have been successfully recovered by using the IPINN, and the network achieved relative L_2 error of $3.575306e^{-3}$ for the RW $q_1(x,t)$ and

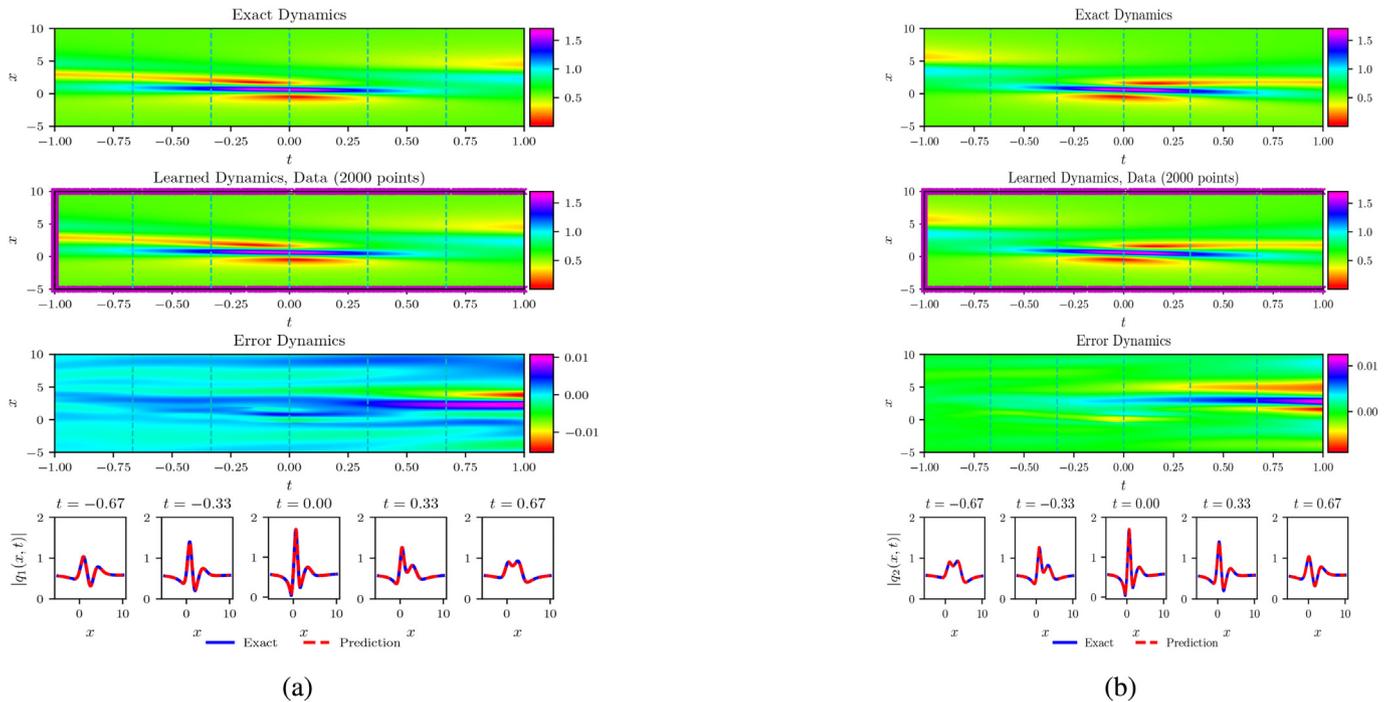


Fig. 14. The vector RWs II $q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediomorchid “ \times ” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector RWs II $q_r(x,t)$ with five distinct training moments $t = -0.67, -0.33, 0.00, 0.33$ and 0.67 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector RWs II $q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the RW II $q_1(x,t)$; (b) The density plots and sectional drawings for the RW II $q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

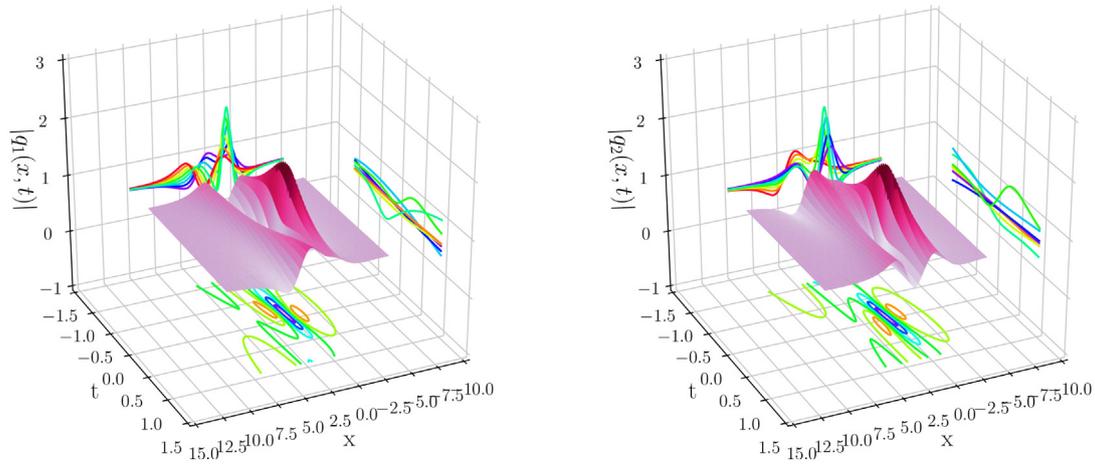


Fig. 15. The three-dimensional plots with contour map on three planes of the predicted vector RWs II $q_r(x,t)$ ($r = 1,2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x,t)$; (Right side panel) The 3D plot for the $q_2(x,t)$.

relative \mathbb{L}_2 error of $3.321024e^{-3}$ for the RW $q_2(x,t)$, and the total number of iterations is 43,068.

Figs. 14–16 exhibit the corresponding training results stemmed from the IPINN for the vector RWs II $q_r(x,t)$ ($r = 1,2$) of the Manakov system with the initial boundary value problem (30) and (31). In Fig. 14, the exact, learned and error dynamics density plots with corresponding amplitude scale size on the right side have been given out, it is worth mentioning that the $N_q = 2000$ training data points involved in the initial-boundary condition are marked by mediumorcid symbol “ \times ” in the learned density graphs both in (a) and (b) of Fig. 14. Meanwhile, the sectional drawings which include the learned and exact vector RWs II $q_r(x,t)$ have been shown at the five distinct moments pointed out in the exact, learned and error dynamics density plots by using darkturquoise dashed lines in the bottom panel of Fig. 14. Similar to the predicted vector RWs I, from the density plots of learning dynamics and section diagrams in Fig. 14, one can see that the two predicted component solutions $q_r(x,t)$ are mirror symmetrical about the central axis of panels (a) and (b) of Fig. 14. Fig. 15 displays the three-dimensional plots with contour map on three planes of the predicted vector RWs II $q_r(x,t)$ based on the IPINN. Fig. 16 exhibits the loss function curve figures of the vector RWs II $q_r(x,t)$ arising from the IPINN with the 20,000 steps Adam and 23,068 steps L-BFGS optimizations on the loss function $\mathcal{L}(\bar{\Theta})$.

4.3. Data-driven vector rogue waves III

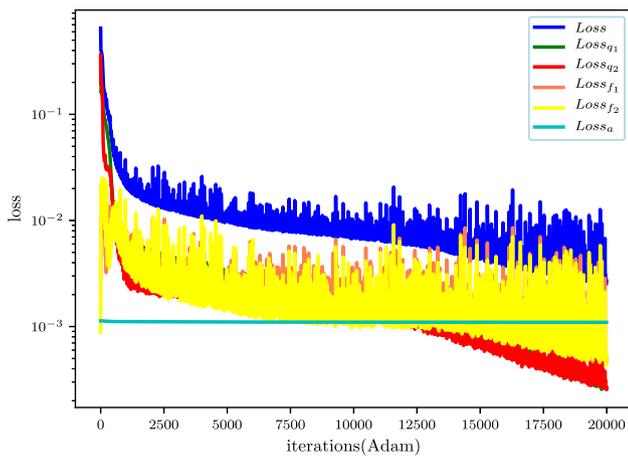
In order to obtain the data-driven vector RWs III of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ by using the 9-layer IPINN with 40 neurons per layer, we are committed to studying the initial value conditions

$$q_r^0(x) = q_{r,rw3}(x, -0.5), x \in [-8.0, 8.0], \tag{32}$$

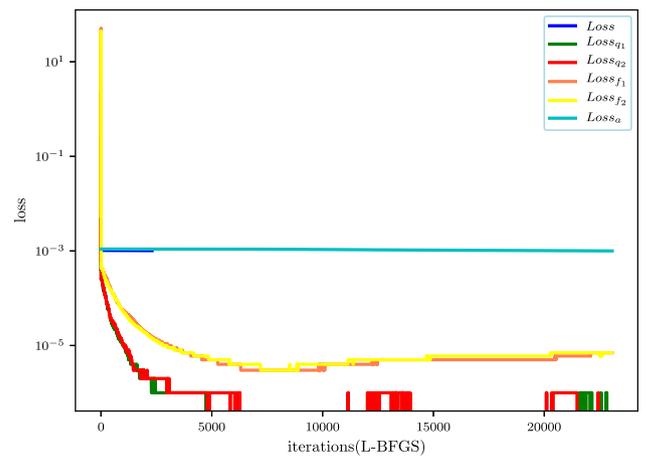
and the Dirichlet boundary conditions

$$q_r^{lb}(t) = q_{r,rw3}(-8.0, t), q_r^{ub}(t) = q_{r,rw3}(8.0, t), t \in [-0.5, 0.5], r = 1, 2. \tag{33}$$

We employ the traditional finite difference scheme on even grids in Matlab to simulate Eq. (25) which contains the initial data (32) and boundary data (33) to obtain the original training data. Specifically, divide spatial region $[-8.0, 8.0]$ into 2000 points and temporal region $[-0.5, 0.5]$ into 1000 points, the vector RWs III are discretized into 1000 snapshots accordingly. We generate a smaller training dataset containing initial-boundary data by randomly extracting $N_q = 2000$ from original dataset and $N_f = 30000$ collocation points which are generated in corresponding spatial-temporal region by utilizing the LHS method. After that, introducing the dataset of initial and boundary



(a)



(b)

Fig. 16. The loss function curve figures of the vector RWs II $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 23,068 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 23,068 L-BFGS optimization iterations.

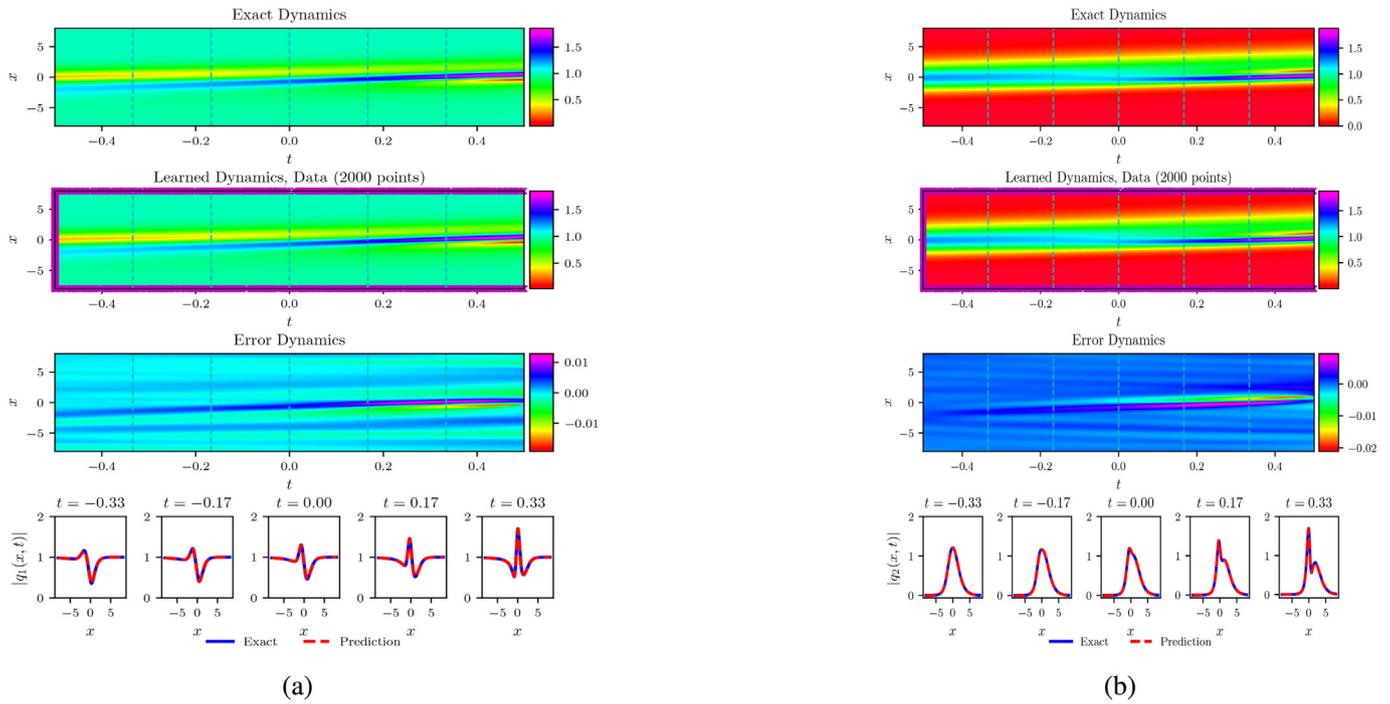


Fig. 17. The vector RWs III $q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “ \times ” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the RWs III $q_r(x,t)$ with five distinct training moments $t = -0.33, -0.17, 0.00, 0.17$ and 0.33 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit the RWs III $q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the RW III $q_1(x,t)$; (b) The density plots and sectional drawings for the RW III $q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

points into the IPINN, and taking 20,000 Adam iterations and 26,037 L-BFGS iterations to optimize loss function $\mathcal{L}(\bar{\theta})$, the latent vector RWs II $q_r(x,t)$ have been successfully learned by tuning all learnable parameters of the IPINN, and the network achieved relative \mathbb{L}_2 error of $2.197789e^{-3}$ for the RW $q_1(x,t)$ and relative \mathbb{L}_2 error of $3.661877e^{-3}$ for the RW $q_2(x,t)$, and the total number of iterations is 46,037.

Figs. 17–19 provide the training results arising from the IPINN for the vector RWs III $q_r(x,t)$ ($r = 1,2$) of the Manakov system with the initial boundary value problem (32) and (33). In Fig. 17, the exact, learned and error dynamics density plots with corresponding amplitude scale size on the right side have been exhibited, it is worth mentioning that the $N_q = 2000$ training data points involved in the initial-boundary

condition are marked by mediumorchid symbol “ \times ” in the learned density plots both in (a) and (b) of Fig. 17. Meanwhile, the sectional drawings which include the learned and exact vector RWs III $q_r(x,t)$ have been shown at the five distinct moments pointed out in the exact, learned and error dynamics density plots by using darkturquoise dashed lines in the bottom panel of Fig. 17. Fig. 18 displays the three-dimensional plots with contour map on three planes of the predicted vector RWs III $q_r(x,t)$ based on the IPINN. From Figs. 17 and 18, with the development of time, one can observe that the first component $q_1(x,t)$ of vector RWs III $q_r(x,t)$ is an interaction solution that dark soliton gradually generates a RW on the side of the dark soliton, while the second component $q_2(x,t)$ is an interaction solution that the bright soliton is gradually emerges a RW on the wave crest of bright soliton.

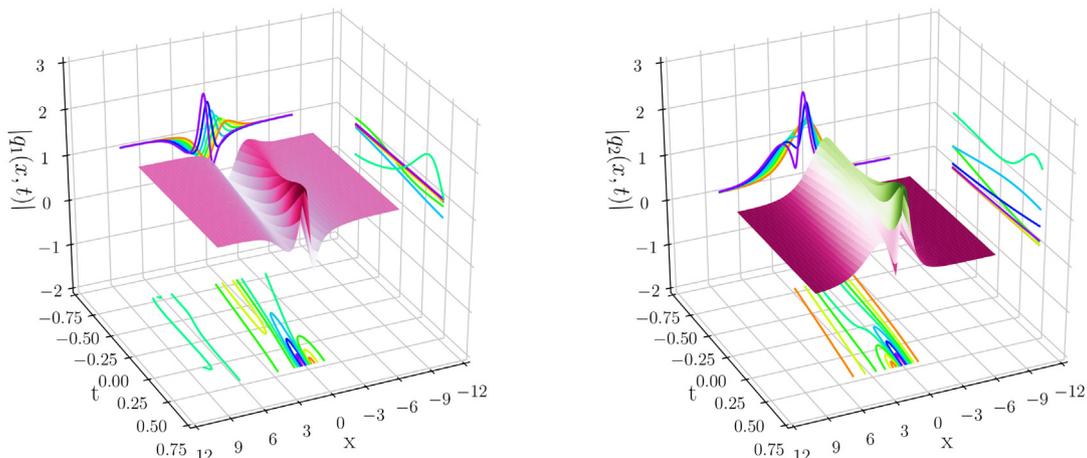


Fig. 18. The three-dimensional plots with contour map on three planes of the predicted vector RWs III $q_r(x,t)$ ($r = 1,2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x,t)$; (Right side panel) The 3D plot for the $q_2(x,t)$.

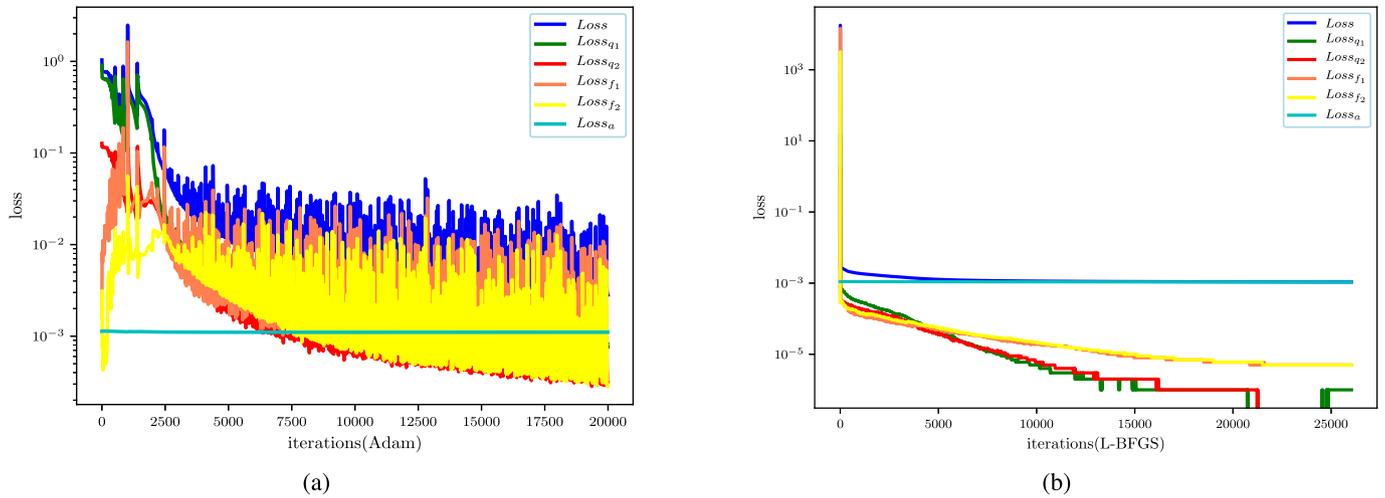


Fig. 19. The loss function curve figures of the vector RWs III $q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 26,037 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 26,037 L-BFGS optimization iterations.

Fig. 19 exhibits the loss function curve figures of the vector RWs III $q_r(x,t)$ arising from the IPINN with the 20,000 steps Adam and 26,037 steps L-BFGS optimizations on the loss function $\mathcal{L}(\bar{\theta})$. In particular, from the (a) of Fig. 19, these loss functions present an unstable iteration curve in the first 2500 iterations of Adam optimizer, and the values of the loss functions have an upward trend at this time.

4.4. Data-driven vector rogue waves IV

Similarly, considering the initial condition and Dirichlet boundary condition of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ to obtain

the data-driven vector RWs 4 by using the 9-layer IPINN with 40 neurons per layer, the $[L_0, L_1]$ and $[T_0, T_1]$ in Eq. (1) are taken as $[-10.0, 5.0]$ and $[-0.5, 0.5]$, respectively. We immediately obtain the initial value conditions

$$q_r^0(x) = q_{r,rw4}(x, -0.5), x \in [-10.0, 5.0], \tag{34}$$

and the Dirichlet boundary conditions

$$q_r^{lb}(t) = q_{r,rw4}(-10.0, t), q_r^{ub}(t) = q_{r,rw4}(5.0, t), t \in [-0.5, 0.5], r = 1, 2. \tag{35}$$

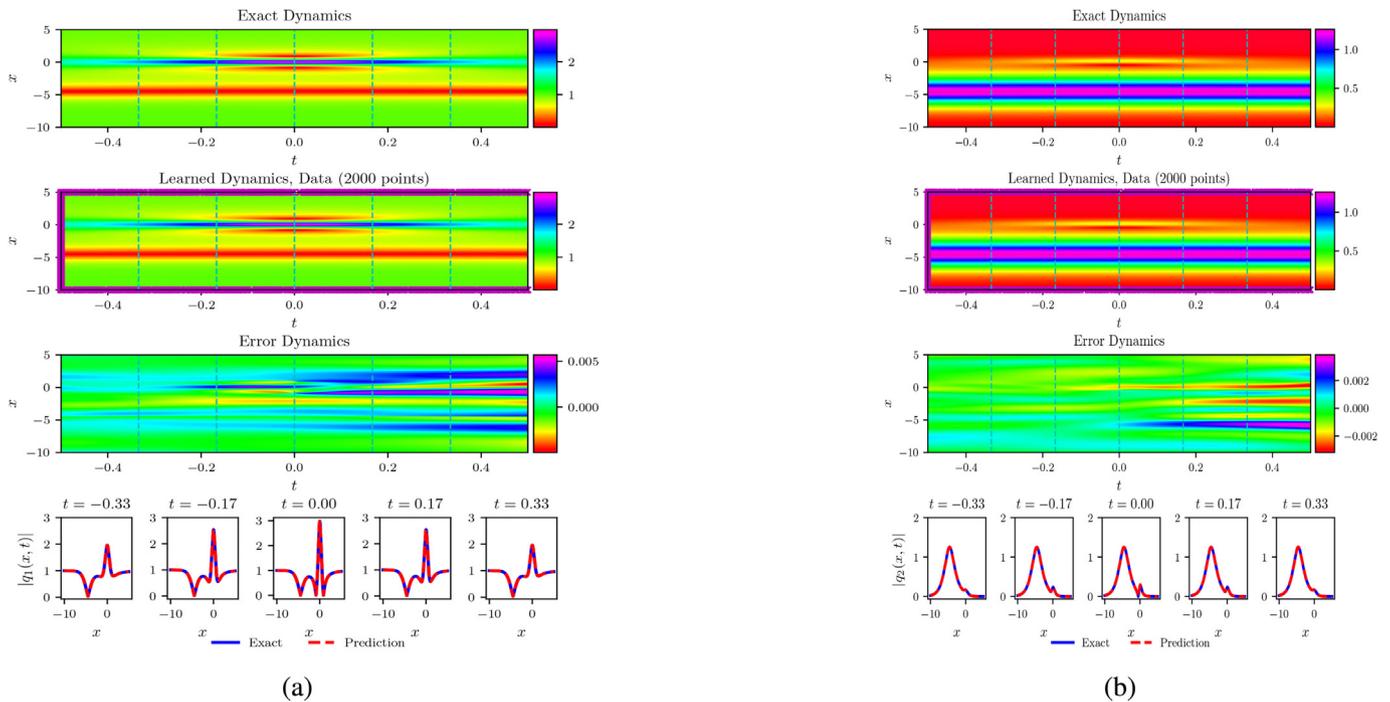


Fig. 20. The vector RWs IV $q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “x” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector RWs IV $q_r(x,t)$ with five distinct training moments $t = -0.33, -0.17, 0.00, 0.17$ and 0.33 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector RWs IV $q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the RW IV $q_1(x,t)$; (b) The density plots and sectional drawings for the RW IV $q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Similarly, discretizing Eq. (26) with the aid of the traditional finite difference scheme on even grids, and we obtain the original training data which contain initial data (34) and boundary data (35) by dividing separately the spatial region $[-10.0, 5.0]$ into 2000 points and the temporal region $[-0.5, 0.5]$ into 1000 points. Then, one can generate a smaller training dataset that contains partial initial-boundary data by randomly extracting $N_q = 2000$ from original dataset and $N_f = 30000$ collocation points which are produced by the LHS. After that, the latent vector RWs $IV q_r(x, t)$ have been successfully learned by tuning all learnable parameters of the IPINN, and the network achieved relative \mathbb{L}_2 error of $1.637892e^{-3}$ for the RW $q_1(x, t)$ and relative \mathbb{L}_2 error of $1.784348e^{-3}$ for the RW $q_2(x, t)$, and the total number of iterations is 51,110.

Figs. 20–22 provide the training results arising from the IPINN for the vector RWs $IV q_r(x, t)$ ($r = 1, 2$) of the Manakov system with the initial boundary value problem (34) and (35). In Fig. 20, the exact, learned and error dynamics density plots with corresponding amplitude scale size on the right side have been exhibited, it is worth mentioning that the $N_q = 2000$ training data points involved in the initial-boundary condition are marked by mediumorchid symbol “ \times ” in the learned density plots both in (a) and (b) of Fig. 20. Meanwhile, the sectional drawings which include the learned and exact vector RWs $IV q_r(x, t)$ have been shown at the five distinct moments pointed out in the exact, learned and error dynamics density plots by using darkturquoise dashed lines in the bottom panel of Fig. 20. Fig. 21 displays the three-dimensional plots with contour map on three planes of the predicted vector RWs $IV q_r(x, t)$ based on the IPINN. From Figs. 20 and 21, with the development of time, one can observe that the first component $q_1(x, t)$ of vector RWs $IV q_r(x, t)$ is the interaction solution composed of dark soliton and RW, while the second component $q_2(x, t)$ is the interaction solution composed of bright soliton and RW whose the amplitude is much lower than that of bright soliton. Different from the case where the RW in the second component of vector RW III emerges on the wave crest, RWs of this part both appear on one side of the bright-dark soliton. Fig. 22 exhibits the loss function curve figures of the vector RWs $IV q_r(x, t)$ arising from the IPINN with the 20,000 steps Adam and 31,110 steps L-BFGS optimizations on the loss function $\mathcal{L}(\bar{\theta})$.

4.5. Data-driven vector rogue waves V

Next, considering the initial conditions $q_r^0(x)$ and Dirichlet boundary conditions $q_r^b(t)$ and $q_r^{ub}(t)$ of Eq. (1) arising from the vector RWs V Eq. (27) for obtaining the data-driven vector RWs 5 by using the 9-layer IPINN with 40 neurons per layer, we set $[L_0, L_1]$ and $[T_0, T_1]$ in

Eq. (1) as $[-8.0, 6.0]$ and $[-0.25, 0.25]$, respectively. After that, the corresponding initial conditions can be written as belows

$$q_r^0(x) = q_{r,rw5}(x, -0.25), x \in [-8.0, 6.0], \tag{36}$$

and the Dirichlet boundary conditions

$$q_r^b(t) = q_{r,rw5}(-8.0, t), q_r^{ub}(t) = q_{r,rw5}(6.0, t), t \in [-0.25, 0.25], r = 1, 2. \tag{37}$$

Here, employing the same data discretization method in Section 4.1, and producing the training data which consists of initial data (36) and boundary data (37) by dividing the spatial region $[-8.0, 6.0]$ into 2000 points and the temporal region $[-0.25, 0.25]$ into 1000 points. We generate a smaller training dataset that containing initial-boundary data by randomly extracting $N_q = 2000$ from original dataset and $N_f = 30000$ collocation points which are generated by the LHS method. After giving the dataset of initial and boundary points, the latent data-driven vector RWs 5 have been successfully learned by tuning all learnable parameters of the IPINN and utilizing 20,000 Adam iterations and 14,602 L-BFGS iterations to regulate the loss function (7), and the network achieved relative \mathbb{L}_2 error of $6.325696e^{-3}$ for the RW $q_1(x, t)$ and relative \mathbb{L}_2 error of $7.654479e^{-3}$ for the RW $q_2(x, t)$, and the total number of iterations is 34,602.

Figs. 23–25 display the deep learning results come from the IPINN for the vector RWs V $q_r(x, t)$ ($r = 1, 2$) of the Manakov system with the initial-boundary value problem (36) and (37). In Fig. 23, the exact, learned and error dynamics density plots with corresponding amplitude scale size on the right side have been exhibited, it is worth mentioning that the $N_q = 2000$ training data points involved in the initial-boundary condition are marked by mediumorchid symbol “ \times ” in the learned density plots both in (a) and (b) of Fig. 23. Meanwhile, the sectional drawings which include the predictive and exact vector RWs V $q_r(x, t)$ have been shown at the five distinct moments pointed out in the exact, learned and error dynamics density plots by using darkturquoise dashed lines in the bottom panel of Fig. 23. Fig. 24 displays the three-dimensional plots with contour map on three planes of the predicted vector RWs V $q_r(x, t)$ based on the IPINN. Fig. 25 exhibits the loss function curve figures of the vector RWs V $q_r(x, t)$ arising from the IPINN with the 20,000 steps Adam and 14,602 steps L-BFGS optimizations on the loss function $\mathcal{L}(\bar{\theta})$.

In this section, we construct a variety of vector RWs of the Manakov system with $\lambda_1 = 1$ and $\lambda_2 = 2$ by using the IPINN, in which contain simple vector RWs in Subsection 4.1 and complex vector RWs in Subsections 4.2–4.5. Although we can accurately simulate both simple

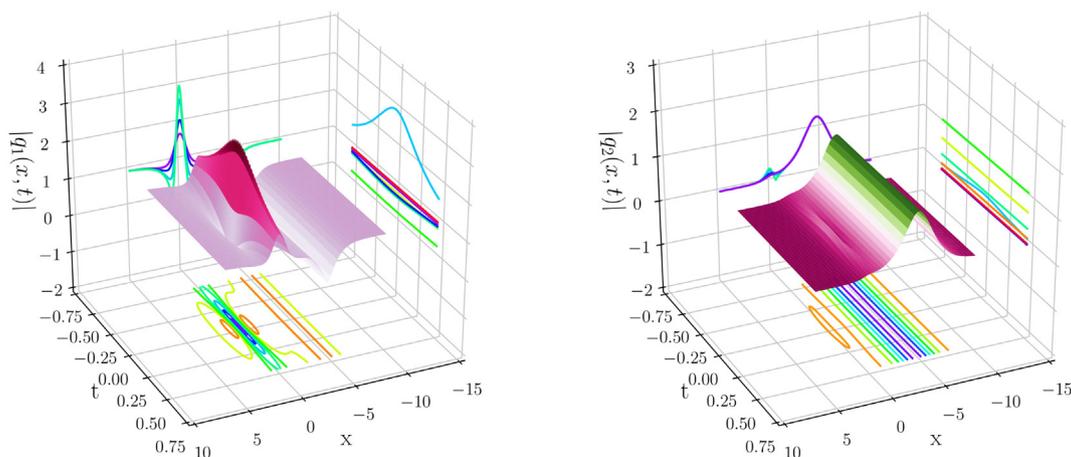


Fig. 21. The three-dimensional plots with contour map on three planes of the predicted vector RWs $IV q_r(x, t)$ ($r = 1, 2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x, t)$; (Right side panel) The 3D plot for the $q_2(x, t)$.

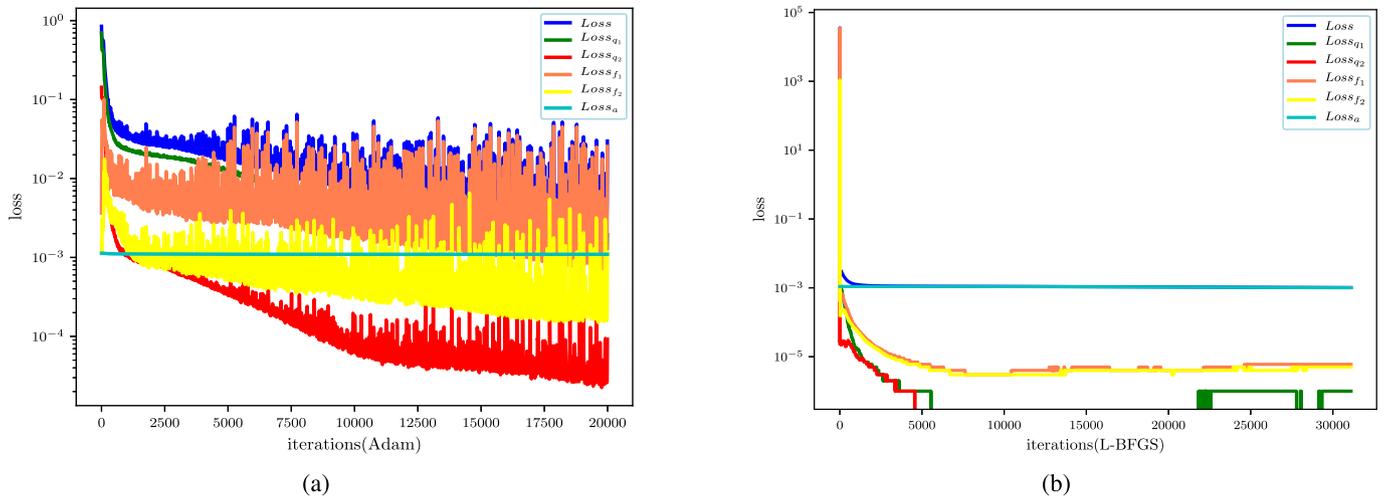


Fig. 22. The loss function curve figures of the vector RWs $V q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 31,110 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 31,110 L-BFGS optimization iterations.

and complex vector RWs, we can only restore their dynamic behavior for complex vector RWs in a short time interval, and a large number of training experiments show that the IPINN has poor dynamic behavior training effect in a long time region. For example, for the vector RWs V formed by the interaction between breathers and RWs, we know that the periodic oscillation can be better displayed in the long-time region. Due to the dynamic structures and initial boundary value condition of complex vector RWs, so we can only capture accurate numerical results in a short time interval. Through a large number of numerical experiments, we are committed to observing the local effect of complex vector RWs more clearly in a long time interval, but we can't get ideal numerical results, so the local effect of complex vector RWs is not obvious. The

main reasons may be as follows: the performance of numerical experimental equipment is insufficient, thus the training intensity of complex systems and their solutions cannot be achieved; the structure of the simulation object itself is complex, we all know that the complex vector RWs changes rapidly with time, unlike the soliton, which does not change with time, so it is difficult to realize the numerical simulation of complex vector RWs in a long time interval; the PINN method itself needs further research to meet the training requirements of complex systems and their numerical solutions. Therefore, how to make the IPINN algorithm train localized waves with good results in the long-time region is an unsolved problem. Finally, we provide a summary of all the aforementioned training results in following Table 2.

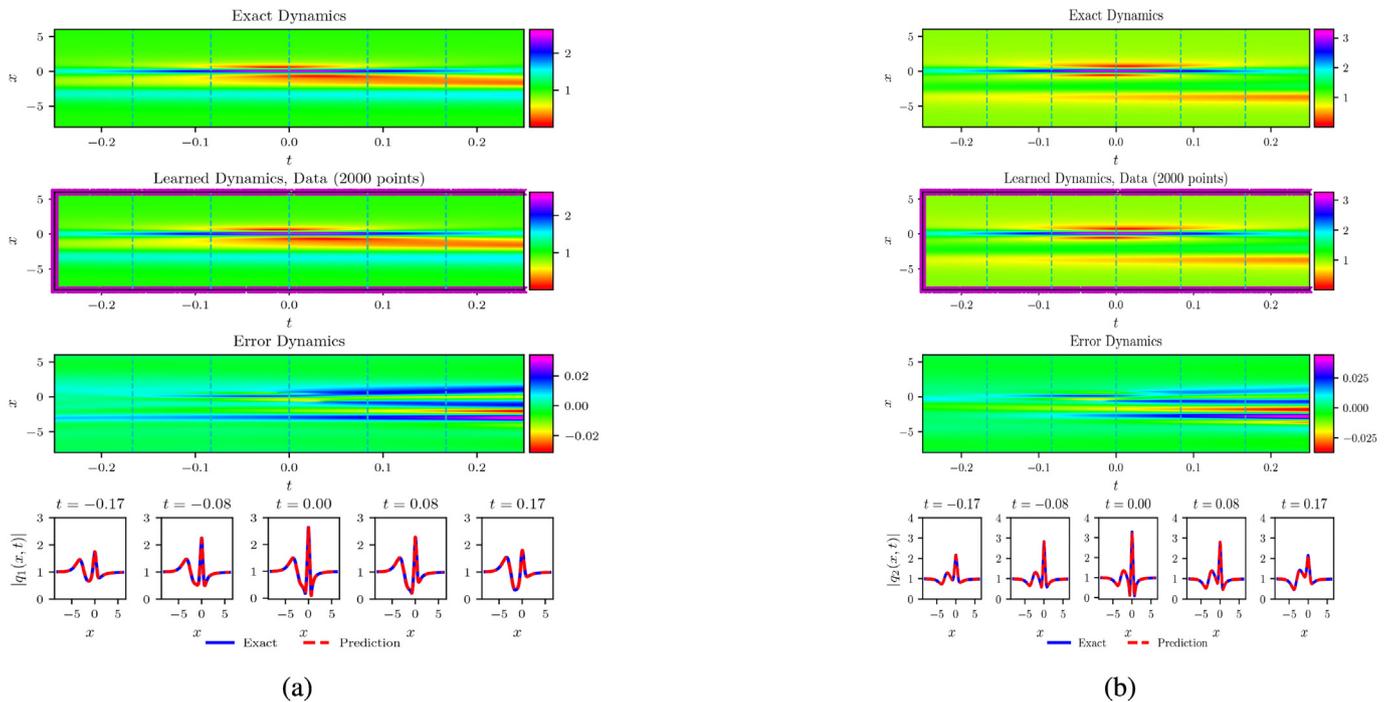


Fig. 23. The vector RWs $V q_r(x,t)$ ($r = 1,2$) resulted from the IPINN with the randomly chosen initial and boundary points $N_q = 2000$ which have been shown by using mediumorchid “ \times ” in learned dynamics, and $N_f = 30000$ collocation points in the corresponding spatiotemporal region. The exact, learned and error dynamics density plots for the vector RWs $V q_r(x,t)$ with five distinct training moments $t = -0.17, -0.08, 0.00, 0.08$ and 0.17 (darkturquoise dashed lines), and the sectional drawings which contain the learned and explicit vector RW $V q_r(x,t)$ at the aforementioned five distinct moments: (a) The density plots and sectional drawings for the RW $V q_1(x,t)$; (b) The density plots and sectional drawings for the RW $V q_2(x,t)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

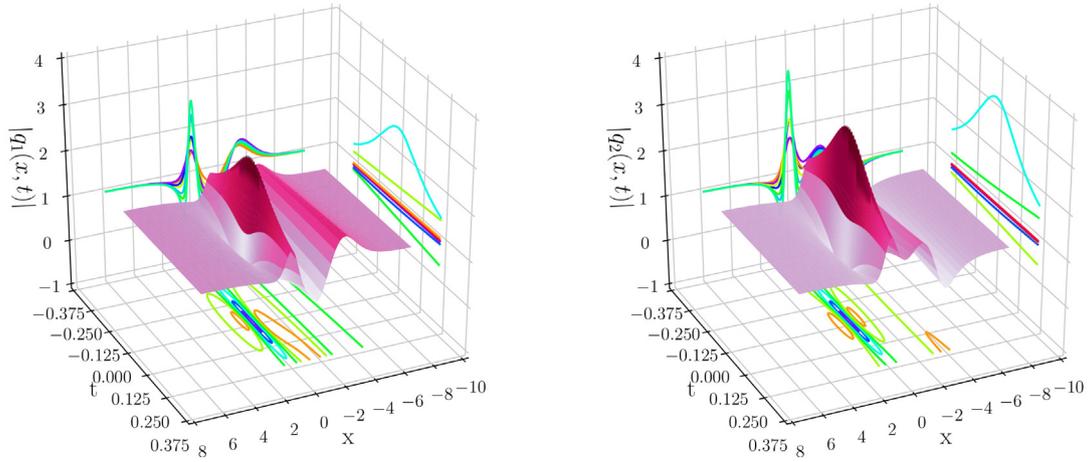


Fig. 24. The three-dimensional plots with contour map on three planes of the predicted vector RWs $V q_r(x,t)$ ($r = 1,2$) based on the IPINN: (Left side panel) The 3D plot for the $q_1(x,t)$; (Right side panel) The 3D plot for the $q_2(x,t)$.

5. Data-driven parameters discovery of Manakov system

In this section, we focus on the data-driven parameters discovery of Manakov system with unknown parameters (1) by using IPINN model. At this time, the physics-informed part becomes the following form

$$\begin{aligned}
 f_u &:= -v_t + \lambda_1 u_{xx} + \lambda_2 u(u^2 + v^2 + m^2 + n^2), \\
 f_v &:= u_t + \lambda_1 v_{xx} + \lambda_2 v(u^2 + v^2 + m^2 + n^2), \\
 f_m &:= -n_t + \lambda_1 m_{xx} + \lambda_2 m(u^2 + v^2 + m^2 + n^2), \\
 f_n &:= m_t + \lambda_1 n_{xx} + \lambda_2 n(u^2 + v^2 + m^2 + n^2),
 \end{aligned}
 \tag{38}$$

in which λ_1 and λ_2 are parameters to be learned. The unknown parameters λ_1 and λ_2 are initialized to $\lambda_1 = \lambda_2 = 0$ in IPINN. Then we define the relative error of unknown parameters is

$$RE = \frac{|\hat{\lambda}_\alpha - \lambda_\alpha|}{\lambda_\alpha} \times 100\%, \alpha = 1, 2,
 \tag{39}$$

where the $\hat{\lambda}_\alpha$ and λ_α represent predicted value and true value, respectively. All noise interference in this section is added to the randomly chosen small data set, the specific form is as follows

$$\begin{aligned}
 Data_train &= Data_train + noise * np.std(Data_train) * np.random. \\
 &randn(Data_train.shape[0], Data_train.shape[1]),
 \end{aligned}
 \tag{40}$$

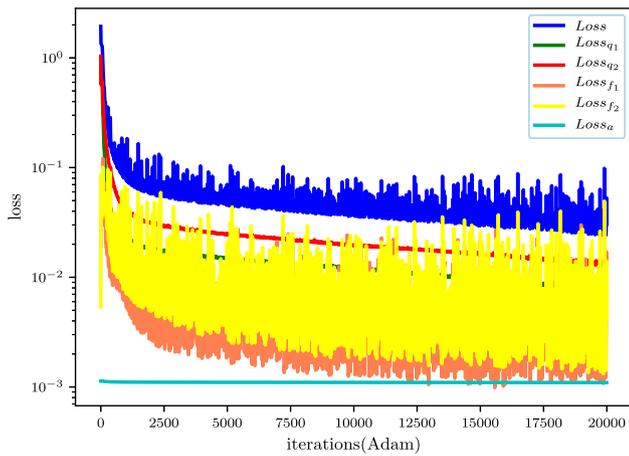
where *noise* and *Data_train* indicate the noise intensity and a small randomly chosen training data set, respectively. The *np.std(·)* returns the standard deviation of an array element, and *np.random.randn(·, ·)* returns a set of samples with a standard normal distribution.

For learning the parameters λ_1 and λ_2 in Eq. (1) with the aid of the IPINN with neuron-wise locally adaptive activation function, and considering the initial conditions and Dirichlet boundary conditions of Eq. (1) arising from the vector RWs I (23) by using the 9-layer IPINN with 40 neurons per layer, we set the spatiotemporal region $(x, t) \in [-3, 6] \times [-0.5, 0.5]$. Thus the corresponding initial conditions can be written as following

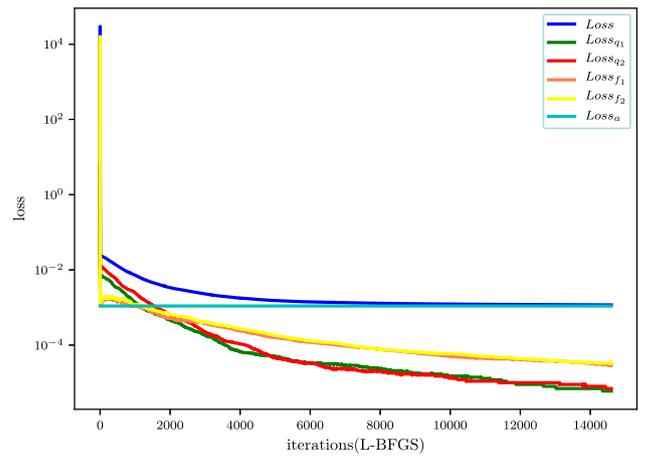
$$q_r^0(x) = q_{r,rw1}(x, -0.5), x \in [-3.0, 6.0],
 \tag{41}$$

and the Dirichlet boundary conditions become

$$q_r^l(t) = q_{r,rw1}(-3.0, t), q_r^{ub}(t) = q_{r,rw1}(6.0, t), t \in [-0.5, 0.5], r = 1, 2.
 \tag{42}$$



(a)



(b)

Fig. 25. The loss function curve figures of the vector RWs $V q_r(x,t)$ ($r = 1,2$) arising from the IPINN with the 20,000 steps Adam and 14,602 steps L-BFGS optimizations: (a) The loss function curve for the 20,000 Adam optimization iterations; (b) The loss function curve for the 14,602 L-BFGS optimization iterations.

Table 2
Relative L_2 errors of three different RWs in IPINN model.

Component	RW Types				
	Vector RWs I	Vector RWs II	Vector RWs III	Vector RWs IV	Vector RWs V
Component q_1	7.505391e-03	3.575306e-03	2.197789e-03	1.637892e-03	6.325696e-03
Component q_2	7.988676e-03	3.321024e-03	3.661877e-03	1.784348e-03	7.654479e-03

Then we use the same data discretization method in Section 4, and generate the original data set by dividing the spatial region $[-3.0,6.0]$ into 2000 points and temporal region $[-0.5,0.5]$ into 1500 points. We obtain a smaller training dataset that containing initial-boundary data (41) and (42) by randomly extracting $N_q = 3000$ from original dataset and $N_f = 30000$ collocation points which are yielded by the LHS method. After that, the latent data-driven unknown parameters λ_1 and λ_2 have been successfully learned by tuning all learnable parameters of the IPINN and utilizing Adam iterations and L-BFGS iterations to regulate the loss function $\mathcal{L}(\bar{\Theta})$.

Fig. 26 displays the training results of unknown parameters under the above initial boundary value conditions by using the IPINN with 20,000 Adam iterations and diverse number of L-BFGS iterations. Fig. 26(a)–(b) exhibit the variation curves of unknown coefficients

λ_1 and λ_2 with different noise intensity. The panel (c) of Fig. 26 shows the variation curve of loss function with different noise intensity. While the error variation plots of unknown coefficients λ_1 and λ_2 under different interference noise are revealed in Fig. 26(d). However, we find that the learning effect of unknown parameters is not ideal, from training result by using clean initial-boundary data (noise=0%) in IPINN, the relative error of λ_1 reaches about 15%, while the relative error of λ_2 exceeds 30% from Fig. 26. Once the noise intensity is increased, the relative error also increases in direct proportion from Fig. 26(d). Thus we urgently need to find IPINN algorithm which can accurately and effectively learn unknown parameters. Fortunately, we find that the PINN with with neuron-wise locally adaptive activation function and L^2 norm parameter regularization shows amazing effect in studying the inverse problem of Yajima-Oikawa system

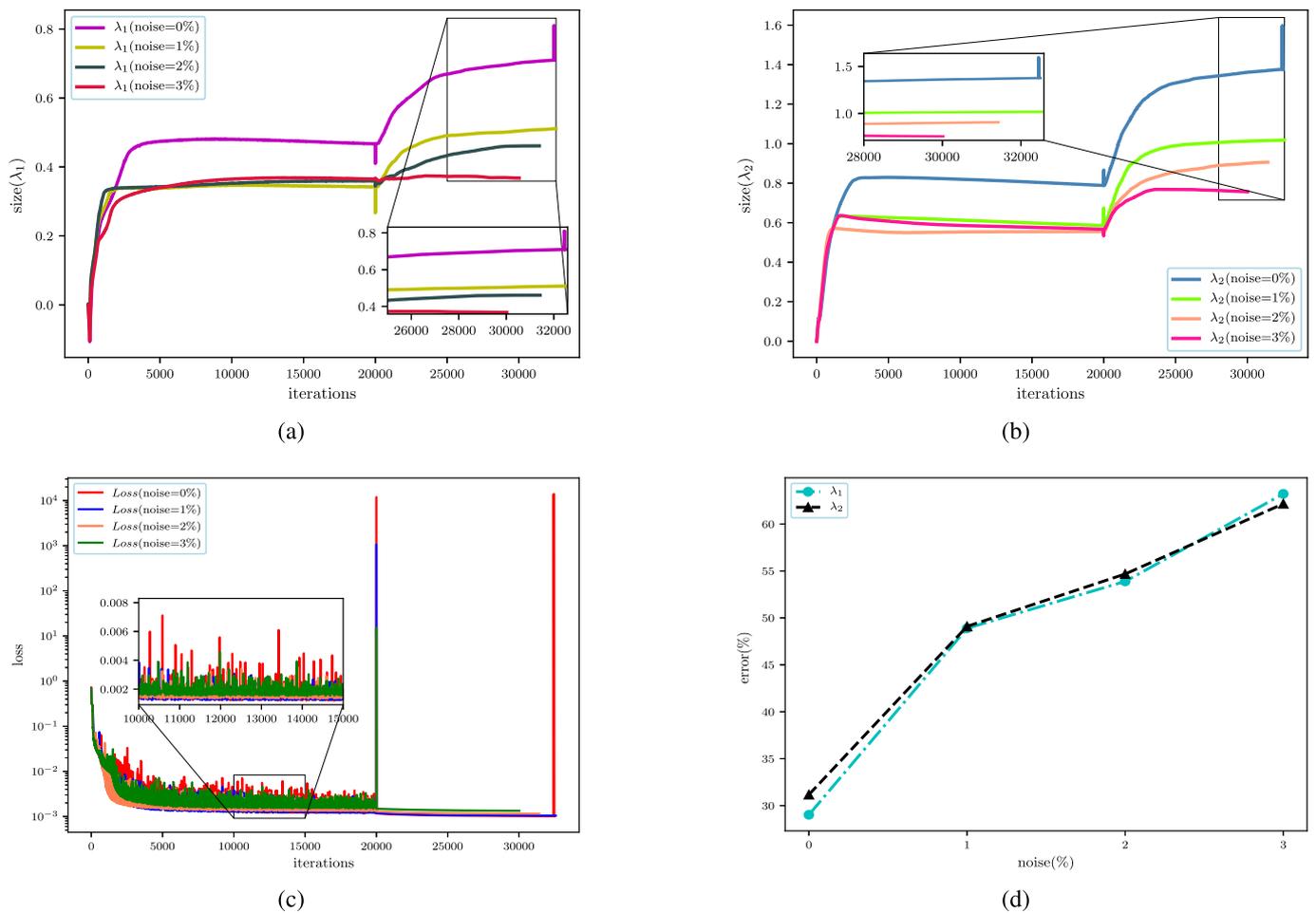


Fig. 26. The training results of parameter discover by means of the IPINN with 20,000 Adam iterations and various number of L-BFGS iterations: (a)–(b) the variation curves of unknown coefficients λ_1 and λ_2 with different noise intensity; (c) the variation curve of loss function with different noise intensity; (d) unknown coefficients λ_1 and λ_2 error variation plots under different interference noise.

[52]. Therefore, we introduce L^2 norm parameter regularization into IPINN to study the parameter discovery problem of Manakov system.

From Ref. [52], we construct a new loss function $\tilde{\mathcal{L}}(\bar{\theta})$ with L^2 norm weight decay, as following

$$\tilde{\mathcal{L}}(\bar{\theta}) = Loss_{PR} = \mathcal{L}(\bar{\theta}) + \frac{\beta}{2} \|W\|_2^2, \quad (43)$$

where $\mathcal{L}(\bar{\theta})$ and W have been defined in Eq. (7) and Eq. (5).

Next, we utilize the same initial-boundary value data points as IPINN without parametric regularization, after 20,000 Adam iterations and various number of L-BFGS iterations, then corresponding numerical results are exhibited in Fig. 27 by using IPINN with $\beta = 0.0001$ weight decay. Fig. 27(a) and (b) display the variation curves of unknown coefficients λ_1 and λ_2 with different noise intensity, Fig. 27(c) depicts the variation curve of loss function with different noise intensity, and the noise intensity and relative error plots are shown in Fig. 27(d). The specific numerical results show that λ_1 obtains the minimum relative error of 2.263904% as the noise intensity is 0%, and maximum relative error 11.419546% as the noise intensity is 1%, whereas λ_2 obtains the minimum relative error of 0.225925% as the noise intensity is 0%, and maximum relative error 13.228714% as the noise intensity is 1%. Compared with the numerical results obtained by IPINN without parameter regularization in Fig. 26, the relative error of parameters learned by IPINN with L^2 norm parameter regularization is small as a whole, especially when the noise intensity is 0%, the parameters

training effect is the best. In other words, the training effect has been very good in the case of clean data. However, when using the data with noise interference for training, the relative error of the prediction parameters is larger than that of the training results of clean data. Therefore, we need to reset the weight decay coefficient to improve the training effect of using the data with noise interference.

Now we attempt to increase the weight decay coefficient by an order of magnitude to further obtain a better training effect in the case of noise data, so we set $\beta = 0.001$ in IPINN with L^2 norm penalty, and employ same initial-boundary value data points as IPINN without parametric regularization, then corresponding training results are depicted in Fig. 28 by using 20,000 Adam iterations and various number of L-BFGS iterations in IPINN with parametric regularization. From Fig. 28, one can find that the training effect with data of various noise is much better than that with clean data, except that the training effect is poor when the noise intensity is 1%. Although the training effect is not ideal when using individual noise intensity data, the generalization training shows that IPINN with coefficient $\beta = 0.001$ of weight decay has excellent noise resistance. Of course, we can also adjust coefficient of the weight decay to adapt to the various problems at hand.

So to summarise, in this section, the parameter discovery problem of Manakov system is studied firstly through IPINN, and it is found that the training effect is not good when using clean data and noisy data. Therefore, we introduce the parameter regularization strategy into IPINN, and study the parameters discovery problem of Manakov system by means of IPINN with two different parameter regularization coefficients.

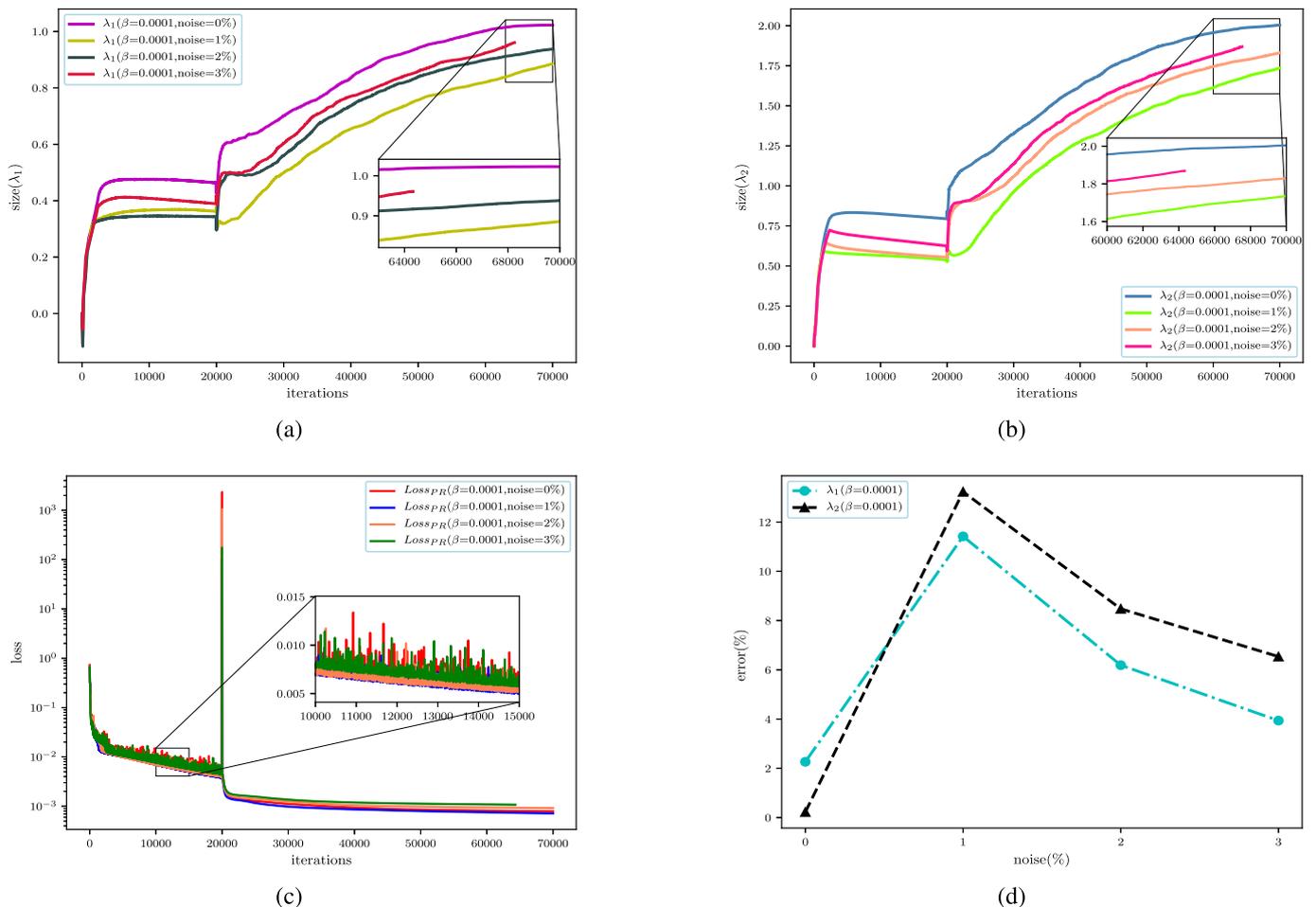


Fig. 27. The training results of parameter discover by means of the IPINN with $\beta = 0.0001$ weight decay: (a)–(b) the variation curves of unknown coefficients λ_1 and λ_2 with different noise intensity; (c) the variation curve of loss function with different noise intensity; (d) unknown coefficients λ_1 and λ_2 error variation plots under different interference noise.

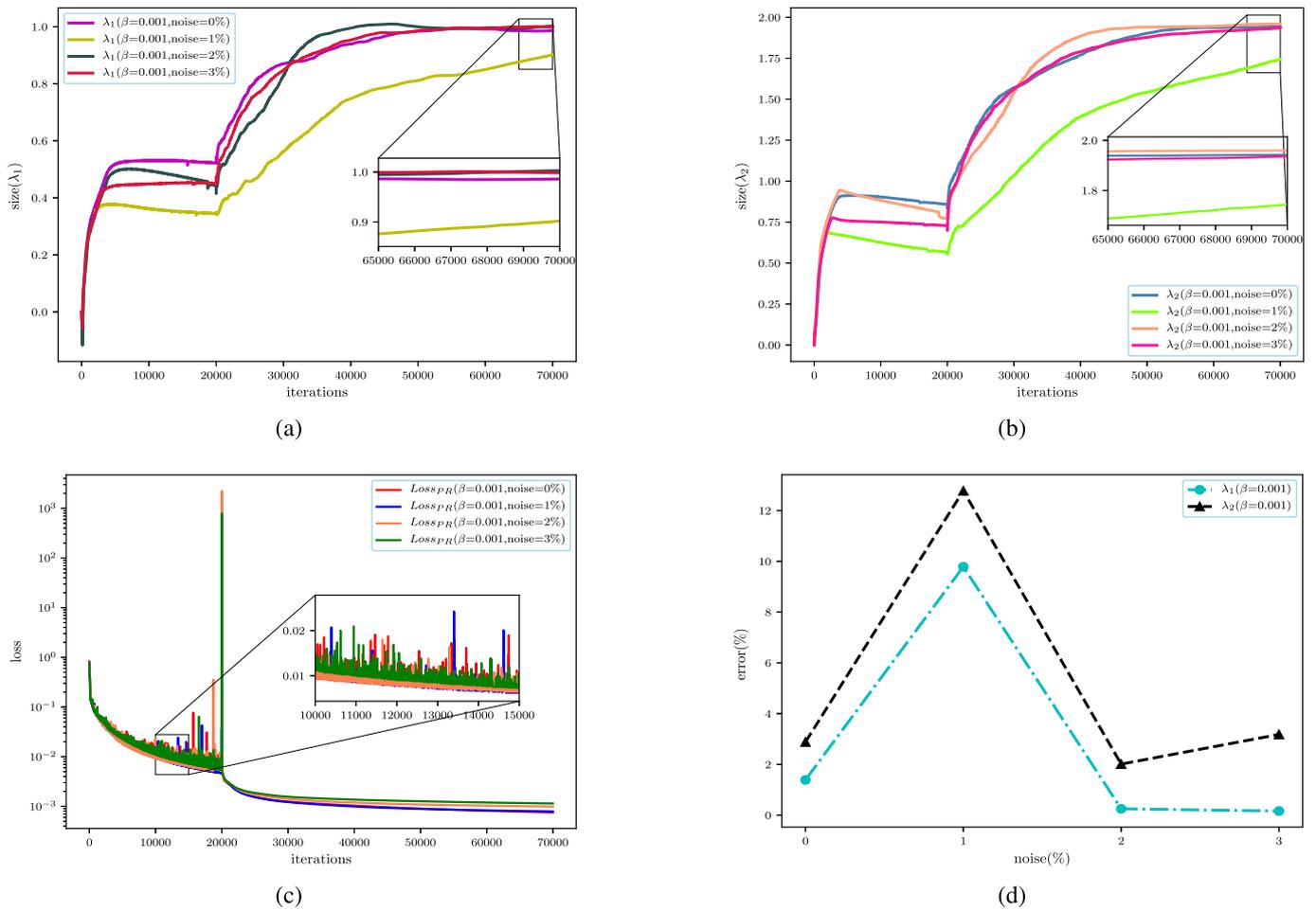


Fig. 28. The training results of parameter discover by means of the IPINN with $\beta = 0.001$ weight decay: (a)–(b) the variation curves of unknown coefficients λ_1 and λ_2 with different noise intensity; (c) the variation curve of loss function with different noise intensity; (d) unknown coefficients λ_1 and λ_2 error variation plots under different interference noise.

Specifically, we find that when the weight decay coefficient $\beta = 0.0001$, the training effect is excellent when using clean data, while the numerical results are not good when using noisy data. However, when the weight decay coefficient $\beta = 0.001$, not only the training effect using noisy data is better than that using clean data, but also the relative error of parameters is small as a whole. Finally, we provide a summary of all the aforementioned training results in following Table 3.

6. Conclusions and discussions

In this paper, we propose an IPINN framework with neuron-wise locally adaptive activation function and slope recovery term for solving data-driven localized waves and recovering unknown parameters of Manakov system for the first time, which provides a very important theoretical basis and training experience for investigating data-driven localized waves and parameters discovery of other two-component or even multi-component coupled nonlinear systems. Since the Manakov system possesses two component solutions $q_r(x, t)$ ($r = 1, 2$), the number of initial-boundary value conditions is twice that of the $(1 + 1)$ dimensional single NLS [42], thus the IPINN framework yield four outputs and four nonlinear equation constraints instead of two outputs and two nonlinear equation constraints of the $(1 + 1)$ dimensional single NLS, so we extended and improved PINN algorithm by increasing the number of output components and physics constraints. Then we are committed to study the data-driven vector localized waves, which contain vector solitons, vector breathers and various vector RWs, as well as parameters discovery for the Manakov system by employing the IPINN

approach with small sample data set. The abundant numerical results show that the IPINN model can effectively and accurately recover the different dynamical behaviors of localized waves for Manakov system fairly. However, we firstly study the parameters discover of Manakov system with unknown parameters by means of the IPINN, and find that the training effect is very poor whether using clean data or noisy data. Therefore, in order to train the parameters more accurately, we introduce the L^2 norm parameter regularization with adjustable weight coefficients β into the routine IPINN, and find that once using the appropriate weight coefficients, the unknown parameters can be recovered accurately and effectively whether using clean data or noise data.

Compared with classical PINN [42], the PINN method with adaptive activation function and slope recovery term loss function has better training effect and better convergence speed for training derivative NLS by means of a large number of training experiments and vivid plots [46,47]. Therefore, starting from the aforementioned PINN approach, we establish an IPINN for investigating data-driven vector localized waves and parameters discovery of Manakov system. Furthermore, in the original PINN model, we only employ the L-BFGS optimization algorithm to optimize the loss function, but we utilize not only the L-BFGS optimizer, but also the Adam optimizer to optimize the loss function in the novel IPINN framework. Moreover, we believe that we can easily construct the IPINN of more component coupled nonlinear systems to obtain vector localized waves, and we can also build corresponding IPINN of the high-dimensional multi-component coupled nonlinear systems for studying high-dimensional vector localized waves. These are also the contents that we will further study in the future.

Table 3
Comparison of correct Manakov system and identified Manakov system obtained by means of the IPINN with different noise intensities and weight hyper-parameters β .

Manakov system	Hyper-parameters		
	$\beta = 0$	$\beta = 0.0001$	$\beta = 0.001$
Correct Manakov system	$iq_{1t} + q_{1xx} + 2(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + q_{2xx} + 2(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 0\%$ $\lambda_2 \text{ error: } 0\%$	$iq_{1t} + q_{1xx} + 2(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + q_{2xx} + 2(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 0\%$ $\lambda_2 \text{ error: } 0\%$	$iq_{1t} + q_{1xx} + 2(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + q_{2xx} + 2(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 0\%$ $\lambda_2 \text{ error: } 0\%$
Identified Manakov system (clean data)	$iq_{1t} + 0.709768q_{1xx} + 1.376634(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.709768q_{2xx} + 1.376634(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 29.023170\%$ $\lambda_2 \text{ error: } 31.168324\%$	$iq_{1t} + 1.022639q_{1xx} + 2.004519(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 1.022639q_{2xx} + 2.004519(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 2.263904\%$ $\lambda_2 \text{ error: } 0.225925\%$	$iq_{1t} + 0.986093q_{1xx} + 1.942293(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.986093q_{2xx} + 1.942293(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 1.390725\%$ $\lambda_2 \text{ error: } 2.885365\%$
Identified Manakov system (1% noise)	$iq_{1t} + 0.511133q_{1xx} + 1.018215(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.511133q_{2xx} + 1.018215(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 48.886692\%$ $\lambda_2 \text{ error: } 49.089260\%$	$iq_{1t} + 0.885805q_{1xx} + 1.735426(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.885805q_{2xx} + 1.735426(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 11.419546\%$ $\lambda_2 \text{ error: } 13.228714\%$	$iq_{1t} + 0.902120q_{1xx} + 1.744415(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.902120q_{2xx} + 1.744415(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 9.788042\%$ $\lambda_2 \text{ error: } 12.779254\%$
Identified Manakov system (2% noise)	$iq_{1t} + 0.461134q_{1xx} + 0.906474(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.461134q_{2xx} + 0.906474(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 53.886570\%$ $\lambda_2 \text{ error: } 54.676281\%$	$iq_{1t} + 0.938114q_{1xx} + 1.830483(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.938114q_{2xx} + 1.830483(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 6.188589\%$ $\lambda_2 \text{ error: } 8.475840\%$	$iq_{1t} + 1.002510q_{1xx} + 1.959728(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 1.002510q_{2xx} + 1.959728(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 0.251031\%$ $\lambda_2 \text{ error: } 2.013606\%$
Identified Manakov system (3% noise)	$iq_{1t} + 0.367840q_{1xx} + 0.756801(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.367840q_{2xx} + 0.756801(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 63.216011\%$ $\lambda_2 \text{ error: } 62.159954\%$	$iq_{1t} + 0.960615q_{1xx} + 1.869262(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.960615q_{2xx} + 1.869262(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 3.938526\%$ $\lambda_2 \text{ error: } 6.536925\%$	$iq_{1t} + 0.998383q_{1xx} + 1.936382(q_1 ^2 + q_2 ^2)q_1 = 0$ $iq_{2t} + 0.998383q_{2xx} + 1.936382(q_1 ^2 + q_2 ^2)q_2 = 0$ $\lambda_1 \text{ error: } 0.161713\%$ $\lambda_2 \text{ error: } 3.180903\%$

The IPINN method showcases a series of results of various interesting problems in the interdisciplinary field of applied mathematics and computational science, and opens a new path for using NN to recover unknown solutions and correspondingly discover the unknown parametric equations in mathematical physics and deep learning. This also provides an significant theoretical basis and experimental reference for solving some previously unsolvable big data spatial-temporal problems and high-dimensional science. Furthermore, the problem of solving multi-component coupled nonlinear systems occupies an important position in many scientific fields, this paper also provides a very powerful deep learning NN framework for these disciplines to make more professional research. In future research, we will focus on finding and proposing more efficient NN algorithms to study a variety of nonlinear systems, in which how to improve PINN model with integrable system theory is a significant problem that needs further research.

CRediT authorship contribution statement

Juncai Pu: Conceptualization, Methodology, Design, Formal analysis, Software, Validation, Writing – original draft, Writing – review & editing.

Yong Chen: Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (No. 12175069) and Science and Technology Commission of Shanghai Municipality (No. 21JC1402500 and No. 18dz2271000).

References

- [1] Khaykovich L, Schreck F, Ferrari G, Bourdel T, Cubizolles J, Carr LD, Castin Y, Salomon C. Formation of a matter-wave bright soliton. *Science*. 2002;296:1290–3.
- [2] Akhmediev N, Eleonskii VM, Kulagin NE. Exact first-order solutions of the nonlinear Schrödinger equation. *Theor Math Phys*. 1987;72:183–96.
- [3] Ma YC. The perturbed plane-wave solutions of the cubic Schrödinger equation. *Stud Appl Math*. 1979;60:43–58.
- [4] Akhmediev N, Ankiewicz A, Taki M. Waves that appear from nowhere and disappear without a trace. *Phys Lett A*. 2009;373:675–8.
- [5] Kharif C, Pelinovsky E. Physical mechanisms of the rogue wave phenomenon. *Eur J Mech B Fluids*. 2003;22:603–34.
- [6] Draper L. Freak Ocean waves. *Weather*. 1966;21:2–4.
- [7] Peregrine DH. Interaction of water waves and currents. *Adv Appl Mech*. 1976;16:9–117.
- [8] Akhmediev N, Ankiewicz A, Soto-Crespo JM. Rogue waves and rational solutions of the nonlinear Schrödinger equation. *Phys Rev E*. 2009;80:026601.
- [9] Akhmediev N, Pelinovsky E. Editorial-introductory remarks on “discussion and debate: rogue waves-towards a unifying concept?”. *Eur Phys J Special Topics*. 2010; 185:1–4.
- [10] Chabchoub A, Hoffmann NP, Akhmediev N. Rogue wave observation in a water wave tank. *Phys Rev Lett*. 2011;106:204502.
- [11] Charalampidis EG, Cuevas-Maraver J, Frantzeskakis DJ, Kevrekidis PG. Rogue waves in ultracold bosonic seas. *Romanian Rep Phys*. 2018;70:504.
- [12] Wang LH, Yang CH, Stöckmann HJ, Kaplan L, Heller EJ. Freak waves in the linear regime: a microwave study. *Phys Rev Lett*. 2010;104:093901.
- [13] Shats M, Punzmann H, Xia H. Capillary rogue waves. *Phys Rev Lett*. 2010; 104:104503.
- [14] Stenflo L, Shukla PK. Nonlinear acoustic-gravity waves. *J Plasma Phys*. 2009;75: 841–7.
- [15] Moslem WM, Shukla PK, Eliasson B. Surface plasma rogue waves. *Europhys Lett*. 2011;96:25002.
- [16] Peregrine DH. Water waves, nonlinear Schrödinger equations and their solutions. *J Austral Math Soc Ser B*. 1983;25:16–43.
- [17] Höhmann R, Kuhl U, Stöckmann HJ, Kaplan L, Heller EJ. The height of an th-order fundamental rogue wave for the nonlinear Schrödinger equation. *Phys Lett A*. 2017;381:1714–8.
- [18] Kaup DJ, Malomed BA, Tasgal RS. Internal dynamics of a vector soliton in a nonlinear optical fiber. *Phys RevE*. 1993;48:3049–53.
- [19] Yan ZY. Vector financial rogue waves. *Phys Lett A*. 2011;375:4274–9.
- [20] Bludov YV, Konotop VV, Akhmediev N. Vector rogue waves in binary mixtures of Bose-Einstein condensates. *Eur Phys J Special Topics*. 2010;185:169–80.
- [21] Manakov SV. On the theory of two-dimensional stationary selffocusing of electromagnetic waves. *Sov Phys - JETP*. 1974;38:248–53.
- [22] Ling LM, Guo BL, Zhao LC. High-order rogue waves in vector nonlinear Schrödinger equations. *Phys Rev E*. 2014;89 041201(R).

- [23] Guo LJ, Wang LH, Cheng Y, He JS. Higher-order rogue waves and modulation instability of the two-component derivative nonlinear Schrödinger equation. *Commun Nonlinear Sci Numer Simul.* 2019;79:104915.
- [24] Wang BH, Lu PH, Dai CQ, Chen YX. Vector optical soliton and periodic solutions of a coupled fractional nonlinear Schrödinger equation. *Res Phys.* 2020;17:103036.
- [25] Rao JG, He JS, Kanna T, Mihalache D. Nonlocal χ -component nonlinear Schrödinger equations: bright solitons, energy-sharing collisions, and positons. *Phys Rev E.* 2020;102:032201.
- [26] Wang LL, Luan ZT, Zhou Q, Biswas A, Alzahrani AK, Liu WJ. Bright soliton solutions of the $(2+1)$ -dimensional generalized coupled nonlinear Schrödinger equation with the four-wave mixing term. *Nonlinear Dyn.* 2021;104:2613–20.
- [27] Dai CQ, Zhang JF. Controlling effect of vector and scalar crossed double-ma breathers in a partially nonlocal nonlinear medium with a linear potential. *Nonlinear Dyn.* 2020;100:1621–8.
- [28] Mjolsness E, DeCoste D. Machine learning for science: state of the art and future prospects. *Science.* 2001;293:2051–5.
- [29] Bishop CM. *Pattern recognition and machine learning.* New York: Springer Press; 2006.
- [30] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature.* 1986;323:533–6.
- [31] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436–44.
- [32] Sun XD, Wu PC, Hoi SCH. Face detection using deep learning: an improved faster RCNN approach. *Neurocomputing.* 2018;299:42–50.
- [33] Alipanahi B, Delong A, Weirauch MT, Frey BJ. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat Biotechnol.* 2015;33:831–8.
- [34] Shao ZF, Wang LG, Wang ZY, Du W, Wu WJ. Saliency-aware convolution neural network for ship detection in surveillance video. *IEEE Trans Circuits Syst Video Technol.* 2020;30:781–94.
- [35] Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K. Speech recognition using deep neural networks: a systematic review. *IEEE Access.* 2019;7:19143–65.
- [36] Collobert R, Weston J, Bottou G, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. *J Mach Learn Res.* 2011;12:2493–537.
- [37] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys.* 2019;378:686–707.
- [38] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang SF, Yang L. Physics-informed machine learning. *Nat Rev Phys.* 2021;3:422–40.
- [39] Jagtap AD, Kawaguchi K, Karniadakis GE. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *J Comput Phys.* 2020;404:109136.
- [40] Jagtap AD, Kawaguchi K, Karniadakis GE. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proc R Soc A.* 2020;476.
- [41] Li J, Chen Y. A deep learning method for solving third-order nonlinear evolution equations. *Commun Theor Phys.* 2020;72:115003.
- [42] Pu JC, Li J, Chen Y. Soliton, breather and rogue wave solutions for solving the nonlinear Schrödinger equation using a deep learning method with physical constraints. *Chin Phys B.* 2021;30:060202.
- [43] Peng WQ, Pu JC, Chen Y. PINN deep learning for the Chen-lee-liu equation: rogue wave on the periodic background. *Commun Nonlinear Sci Numer Simul.* 2022;105:106067.
- [44] Miao ZW, Chen Y. Physics-informed neural network method in high-dimensional integrable systems. *Mod Phys Lett B.* 2022;36:2150531.
- [45] Lin SN, Chen Y. A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions. *J Comput Phys.* 2022;457:111053.
- [46] Pu JC, Li J, Chen Y. Solving localized wave solutions of the derivative nonlinear Schrödinger equation using an improved PINN method. *Nonlinear Dyn.* 2021;105:1723–39.
- [47] Pu JC, Peng WQ, Chen Y. The data-driven localized wave solutions of the derivative nonlinear Schrödinger equation by using improved PINN approach. *Wave Motion.* 2021;107:102823.
- [48] Fang Y, Wu GZ, Wang YY, Dai CQ. Data-driven femtosecond optical soliton excitations and parameters discovery of the high-order NLSE using the PINN. *Nonlinear Dyn.* 2021;105:603–16.
- [49] Wang L, Yan ZY. Data-driven rogue waves and parameter discovery in the defocusing nonlinear Schrödinger equation with a potential using the PINN deep learning. *Phys Lett A.* 2021;404:127408.
- [50] Wu GZ, Fang Y, Wang YY, Wu GC, Dai CQ. Predicting the dynamic process and model parameters of the vector optical solitons in birefringent fibers via the modified PINN. *Chaos Solitons Fractals.* 2021;152:111393.
- [51] Mo YF, Ling LM, Zeng DL. Data-driven vector soliton solutions of coupled nonlinear Schrödinger equation using a deep learning algorithm. *Phys Lett A.* 2022;421:127739.
- [52] Pu JC, Chen Y. Data-driven forward-inverse problems for Yajima-Oikawa system using deep learning with parameter regularization. *arXiv: 2112.04062; 2021.*
- [53] Kaplan AE. Light-induced nonreciprocity, field invariants, and nonlinear eigenpolarizations. *Opt Lett.* 1983;8:560–2.
- [54] Wang D, Menyuk CR. Polarization evolution due to the Kerr nonlinearity and chromatic dispersion. *J Lightwave Technol.* 1999;17:2520–9.
- [55] Yang JK. *Nonlinear waves in integrable and nonintegrable systems.* SIAM, Philadelphia Press; 2010.
- [56] Vishnu Priya N, Senthilvelan M, Lakshmanan M, Akhmediev breathers, ma solitons, and general breathers from rogue waves: a case study in the manakov system. *Phys Rev E.* 2013;88:022918.
- [57] Guo BL, Ling LM. Rogue wave, breathers and bright-dark-rogue solutions for the coupled Schrödinger equations. *Chin Phys Lett.* 2011;28:110202.
- [58] Baronio F, Degasperis A, Conforti M, Wabnitz S. Solutions of the vector nonlinear Schrödinger equations: evidence for deterministic rogue waves. *Phys Rev Lett.* 2012;109:044102.
- [59] Zhao LC, Liu J. Localized nonlinear waves in a two-mode nonlinear fiber. *J Opt Soc Am B.* 2012;29:3119–27.
- [60] Wang X, Yang B, Chen Y, Yang YQ. Higher-order localized waves in coupled nonlinear Schrödinger equations. *Chin Phys Lett.* 2014;31:090201.
- [61] Rao JG, Porsezian K, Kanna T, Cheng Y, He JS. Vector rogue waves in integrable χ -coupled nonlinear Schrödinger equations. *Phys Scr.* 2019;94:075205.
- [62] Kingma DP, Ba JL. Adam: a method for stochastic optimization. *arXiv: 1412.6980; 2014.*
- [63] Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math Program.* 1989;45:503–28.
- [64] Stein M. Large sample properties of simulations using latin hypercube sampling. *Technometrics.* 1987;29:143–51.