Search for rogue waves in Bose-Einstein condensates via a theory-guided neural network

Xiao-Dong Bai[1] and Dongxiao Zhang[2,*]

[1]*Department of Mathematics and Theories, Peng Cheng Laboratory, Shenzhen 518055, Guangdong, People's Republic of China*

[2]*National Center for Applied Mathematics Shenzhen (NCAMS), Southern University of Science and Technology, Shenzhen 518055, Guangdong, People's Republic of China*

(Received 14 February 2022; accepted 25 July 2022; published 5 August 2022)

An important and incompletely answered question is whether machine learning methods can be used to discover the excitation of rogue waves (RWs) in nonlinear systems, especially their dynamic properties and phase transitions. In this work, a theory-guided neural network (TgNN) is constructed to explore the RWs of one-dimensional Bose-Einstein condensates. We find that such method is superior to the ordinary deep neural network due to theory guidance of underlying problems. The former can directly give any excited location, timing, and structure of RWs using only a small amount of dynamic evolution data as the training data, without the tedious step-by-step iterative calculation process. In addition, based on the TgNN approach, a phase transition boundary is also discovered, which clearly distinguishes the first-order RW phase from the non-RW phase. The results not only greatly reduce computational time for exploring RWs, but also provide a promising technique for discovering phase transitions in parameterized nonlinear systems.

DOI: 10.1103/PhysRevE.106.025305

I. INTRODUCTION

Rogue waves (RWs) have attracted increasing attention over the past decades because they have been universally characterized as mysterious due to their remarkable feature in which they appear suddenly and disappear without a trace [1–3]. Thus far, RWs have been widely found in many nonlinear systems, including oceans, the atmosphere [4], optics [5], plasmas [6], superfluids [7], Bose-Einstein condensates (BECs) [8], and capillary waves [9]. Especially in oceans, the sudden occurrence of such extreme events represents a significant threat to offshore platforms, ships, and renewable energy devices [1]. Therefore, obtaining their excited locations, timing, and structures in these complex nonlinear systems is not only significant for the safety of human life and property in maritime activities, but also an essential task to elucidate the nature of this unusual phenomenon.

Currently, the structures, dynamics, and other properties of RWs have yielded pioneering achievements, both theoretically [10–12] and experimentally [13,14]. For example, in terms of theory, the interesting transverse stability of RWs has recently been studied in the NLS equation [15], and feasible schemes for generating them are also proposed [16]. Moreover, excitations of RWs have been observed experimentally in water tanks [17], nonlinear fiber optics [18], and multifractal photonic arrays [19]. Other important investigations of RWs have also produced notable results [20–29].

Significantly, although RWs have been widely investigated, these results mainly rely on real-time evolution methods (RTEM) with the iterative calculation process. If one intends to predict in advance a possible excited location, timing, and structure of the RW that needs to be explored, substantial time and computing resources will be required. This is especially true in cases in which the system being studied is large and complex, because the tedious step-by-step iterative calculation process is inevitable in current traditional RTEM. Consequently, it is nonpragmatic to find all possible RWs that have not been discovered within a certain parameter range by current methods, and developing a universal approach to replace these traditional methods is urgently needed.

In recent years, machine learning (ML) technique, as a powerful tool for analyzing data, has achieved enormous success, from industrial applications to fundamental research [30–44]. In particular, the study of dynamic characteristics and phase transitions of physical systems has been actively pursued using the deep neural network (DNN) [45–49]. However, two problems are commonly encountered in DNN methods. First, data acquisition is often difficult and time consuming in many scientific researches; and second, the DNN itself does not consider prior theories that it should obey for specific physical problems, leading to some predictions that may violate common sense in physics. In order to overcome these deficiencies, the physics-informed neural network (PINN) has recently been proposed, in which a constraint term from physical laws that it should obey is incorporated into neural network training [50–52]. Furthermore, the theory-guided neural network (TgNN) framework has also been developed, which can not only incorporate physics principles, but also integrate practical engineering theories into neural network training [53]. Adding these theories to guide the training of DNN can not only improve the accuracy of neural network predictions, but also reduce its dependence on data.

Based on these advantages of neural networks and the shortcomings of traditional research methods of RWs, in this paper, we incorporate the Gross-Pitaevskii equations that it

*zhangdx@sustech.edu.cn

should obey into the neural network. Then, a TgNN approach is developed to replace traditional methods to more quickly give any excited location, timing, and structure of RWs in a one-dimensional (1D) BEC, thus avoiding the tedious step-by-step iterative calculation process. Further, based on this method, a phase transition boundary is also discovered, which clearly distinguishes the first-order RW phase from the non-RW phase.

The remainder of this paper is organized as follows. In Sec. II, we introduce the model of the dimensionless 1D BECs and briefly review the structures of first-order RWs. In Sec. III, We introduce the network structure and operation mechanism of DNN and TgNN methods in detail. In Sec. IV, we compare and discuss the results of DNN and TgNN for RW prediction. Finally, we conclude in Sec. V.

II. MODEL

We consider dilute degenerate bosons confined in a cigar-shaped trap and assume that the system is far from the Tonks-Girardeau regime [54]. Then, the BECs can be well described by the dimensionless 1D Gross-Pitaevskii (GP) equation:

$$i\frac{\partial\psi(x,t)}{\partial t} = \left[-\frac{1}{2}\frac{\partial^2}{\partial x^2} + g|\psi(x,t)|^2 + V(x)\right]\psi(x,t), \quad (1)$$

where the coordinate x is measured in units of $\sqrt{\hbar/m\omega_x}$; time is in units of $1/\omega_x$; ω_x is the x component frequency of the harmonic trap; $\psi(x,t)$ is the macroscopic wave function of the condensate, and it satisfies $\int|\psi(x,t)|^2dx = 1$; $g = 4N\pi\alpha_{1d}a_s/\sqrt{\hbar/m\omega_x}$ characterizes the interatomic interaction, and is determined by the s-wave scattering length a_s (here, we focused on attractive BECs for which $a_s < 0$); and $\alpha_{1d} = \int|\varphi(y,z)|^4dydz/(\int|\varphi(y,z)|^2dydz)^{5/2}$ is a coefficient, which compensates for the loss of two dimensions [55] and depends on the transversal directions being integrated out. In the above expressions, m is the mass of atom; N is the total number of atoms; and $\varphi(y,z)$ is the ground wave function of the lateral dimensions. Finally, the potential $V(x) = \frac{1}{2}\Omega^2x^2$, and we here assume $\Omega = \Omega_0 = 0.02$. These settings are easy to implement experimentally [56], and related hyperfine states or isotopes of the observed dynamics could also be experimentally traced.

Next, we assume that the initial wave packet is a simple Gaussian function

$$\psi(x, t=0) = C * \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (2)$$

where σ and C denote the width and normalized coefficient of the wave packet, respectively; and the boundary conditions are $\psi(x=-40) = \psi(x=40) = 0$, and here we have set $x \in (-40, 40)$. It has been demonstrated that, if σ is large, the initial Gaussian wave packet will expand and evolve into a structure of a "Christmas tree" [16,57,58] (i.e., the peak emergence from one to two, then to three, etc.), as shown in Fig. 1(a). The first high-amplitude waveform marked by a white square is called the Peregrine soliton [2], which is the typical first-order RW, and its detailed structure is shown in Fig. 1(b). However, the peak marked by white dotted box D in Fig. 1(a) has the same structure of first-order RW, but its amplitude is lower than that of Peregrine soliton. In order to distinguish them we temporarily name it as weak-first-order

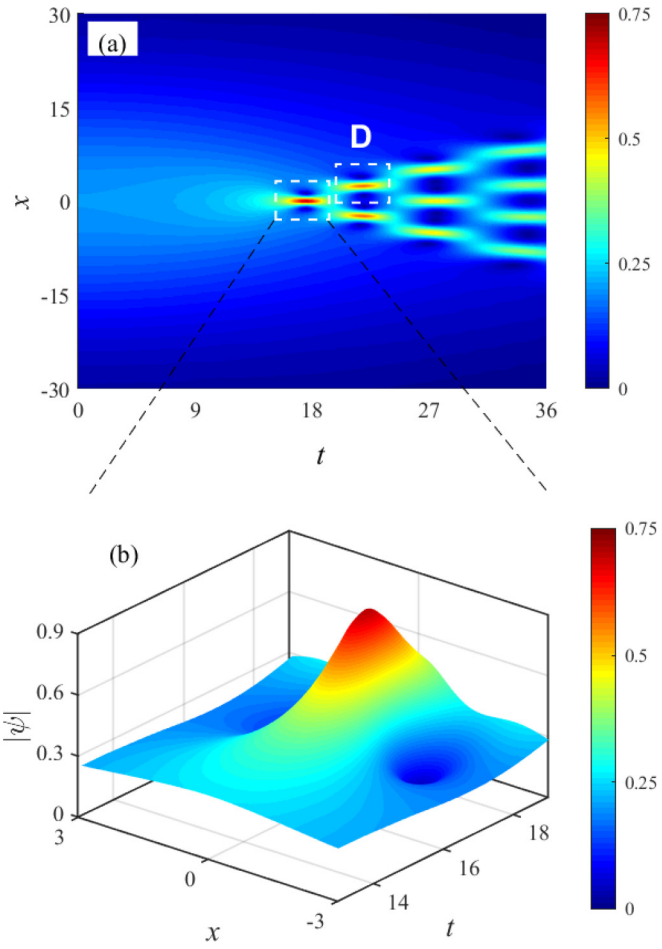

FIG. 1. (a) Dynamics of one-dimensional BECs with the initial condition Eq. (2) as time, which are solved from Eq. (1) numerically. The first excitation marked by a white square is a first-order RW, which is also called the Peregrine soliton. (b) The detailed structure of first-order RW of (a). The color represents $|\psi(x,t)|$ in both (a) and (b). The parameters are: $g = -8$ and $\sigma = 20$, and C is determined by normalization.

RW. It is important to note that this phenomenon does not appear in all parameter cases, but strongly depends on both g and σ. Indeed, that is the subject addressed by this work.

III. METHOD

In general, the structure of DNN contains an input layer, several hidden layers, and an output layer. As a powerful nonlinear function approximation tool, it can learn complex projection mapping between inputs and outputs with a simple network [53,59]. If we assume that there are L hidden layers, the forward formulation of DNN can be described as follows:

$$\mathbf{F}^n = \Theta_n(\mathbf{W}^n\mathbf{X}^{n-1} + \mathbf{b}^n), \quad n = 1, 2, \ldots, L+1, \quad (3)$$

where $\mathbf{X}^0$ and $\mathbf{F}^{L+1}$ are the input and output vectors, respectively; $\mathbf{F}^n$, $n = 1, 2, \ldots, L$, is the output of the nth layer, and is also the input of the $(n+1)$th layer; Θ_n is the activation function of the nth layer, such as rectified linear unit (ReLU), hyperbolic tangent (Tanh), and sigmoid; $\mathbf{W}^n$ and $\mathbf{b}^n$ are the weights and biases matrices of the nth layer, respectively; and

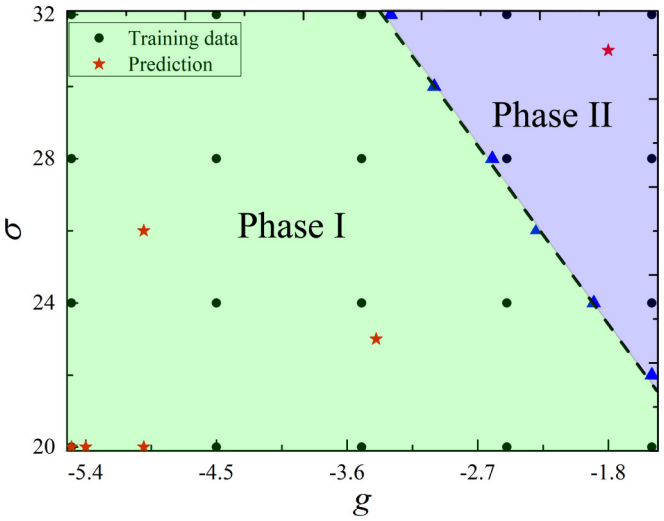

FIG. 2. The black dots represent the training data obtained by the traditional real-time evolution method. The red stars represent the cases that are used to test and compare the performance of the neural network that has been successfully trained, which correspond to Fig. 4 and Fig. 6. Phases I–II correspond to the first-order RW phase and the non-RW phase, respectively. They are clearly distinguished by the black dotted line obtained by the TgNN method. The blue upper triangles represent the same phase transition boundary attained by the traditional method.

the input of the $(n-1)$th layer is $\mathbf{X}^{n-1}$. The loss function, being the mean square error (MSE) between the output of the network and the truth data, can be represented as follows:

$$L_{NN} = \frac{1}{N_{data}} \sum_{k=1}^{N_{data}} \left(\phi_k^{NN}(\mathbf{X}_k; \mathbf{W}, \mathbf{b}) - \phi_k\right)^2, \tag{4}$$

where N_{data} denotes the total number of training data points; and ϕ^{NN} and ϕ are the prediction of the neural network and the true data, respectively. Obviously, the smaller the L_{NN} is, the closer the predicted result is to the truth data. During the training process, the network parameters $(\mathbf{W}, \mathbf{b})$ can be optimized and updated by minimizing the loss function with the algorithms of stochastic gradient descent or Adam. It should be noted that when the neural network is actually trained to replace the process of solving GP equations, the input data should include (g, σ, x, t). Once the neural network is successfully trained, we can directly give the corresponding dynamic process for any parameters without the tedious step-by-step iterative calculation process.

In this work, our aim is to introduce a TgNN model based on DNN to explore the RWs of the system described by Eq. (1). As Fig. 2 shows, at the parameters marked by the black dots, the dynamic processes of forming RWs have been obtained in advance using traditional RTEM, and their entire real-time evolution data have been recorded as training data. Then, our intention is to use the TgNN method, bypassing the iterative calculation process of RTEM, to rapidly discover any unknown excited location, timing, and structure of RWs within this parameter plane, using only these small amounts of known data as the training data. In addition, a phase transition boundary needs to be given based on this method to determine which range of RWs can be excited, and which ones cannot, within our observation period.

Prior to elaborating the structure of TgNN [53], it is necessary to first separate the real and imaginary parts of the wave function $\psi(x,t) = u(x,t) + iv(x,t)$ in Eq. (1), because the neural network does not support complex operations. Thus, Eq. (1) becomes:

$$\begin{aligned} \frac{\partial u}{\partial t} &= -\frac{1}{2}\frac{\partial^2 v}{\partial x^2} + gv(u^2+v^2) + \frac{1}{2}\Omega_0^2 x^2 v \\ \frac{\partial v}{\partial t} &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - gu(u^2+v^2) - \frac{1}{2}\Omega_0^2 x^2 u, \end{aligned} \tag{5}$$

where $u(x,t)$ and $v(x,t)$ denote the real and imaginary parts of $\psi(x,t)$, respectively. Therefore, the atomic density is $|\psi(x,t)| = \sqrt{u(x,t)^2 + v(x,t)^2}$. If we assume $\Gamma^{NN}(\cdot) = \Gamma^{NN}(g, \sigma, x, t; \mathbf{W}, \mathbf{b})$ and Γ can be u or v, which represents the NN approximation of the real and imaginary parts for ψ, respectively, then Eq. (5) can be redefined as follows:

$$\begin{aligned} f_u &= -\frac{\partial u^{NN}(\cdot)}{\partial t} - \frac{1}{2}\frac{\partial^2 v^{NN}(\cdot)}{\partial x^2} + \left(g\rho + \frac{1}{2}\Omega_0^2 x^2\right) v^{NN}(\cdot) \\ f_v &= -\frac{\partial v^{NN}(\cdot)}{\partial t} + \frac{1}{2}\frac{\partial^2 u^{NN}(\cdot)}{\partial x^2} - \left(g\rho + \frac{1}{2}\Omega_0^2 x^2\right) u^{NN}(\cdot), \end{aligned} \tag{6}$$

where $\rho = [u^{NN}(\cdot)]^2 + [v^{NN}(\cdot)]^2$, and the partial derivatives can be easily computed by applying the chain rule for the network through automatic differentiation. In order to enforce the Eq. (5) constraint during the training process, both f_u and f_v should approach zero. The NN approximations of initial and boundary conditions Γ_I and Γ_B can be denoted by $\Gamma_I^{NN}(\cdot)$ and $\Gamma_B^{NN}(\cdot)$, respectively. The loss function of TgNN is then jointly determined by the data mismatch (R_{data}), strong form residual (R_f), and initial and boundary condition regularization (R_I and R_B):

$$\begin{aligned} R_{data}^{\Gamma} &= \frac{1}{N_d}\sum_{k=1}^{N_d}\left[\Gamma_k^{NN}(\cdot) - \Gamma_k(\cdot)\right]^2, \\ R_I^{\Gamma} &= \frac{1}{N_I}\sum_{k=1}^{N_I}[\Gamma_{I,k}^{NN}(\cdot) - \Gamma_{I,k}]^2, \\ R_B^{\Gamma} &= \frac{1}{N_B}\sum_{k=1}^{N_B}[\Gamma_{B,k}^{NN}(\cdot) - \Gamma_{B,k}]^2, \\ R_f^{\Gamma} &= \frac{1}{N_f}\sum_{k=1}^{N_f}[f_{\Gamma,k}(\cdot)]^2, \end{aligned} \tag{7}$$

where N_d denotes the total number of training data points; N_f is the number of collocation points for Eq. (6) residual evaluation; N_I and N_B are the numbers of collocation points for the evaluation of initial and boundary conditions, respectively. R_{data}^{Γ}, R_I^{Γ}, R_B^{Γ}, and R_f^{Γ} are actually the same as Eq. (4). For example, R_{data}^{Γ} represents

$$R_{data}^{u} = \frac{1}{N_d}\sum_{k=1}^{N_d}\left[u_k^{NN}(g, \sigma, x, t; \mathbf{W}, \mathbf{b}) - u_k(g, \sigma, x, t; \mathbf{W}, \mathbf{b})\right]^2, \tag{8}$$

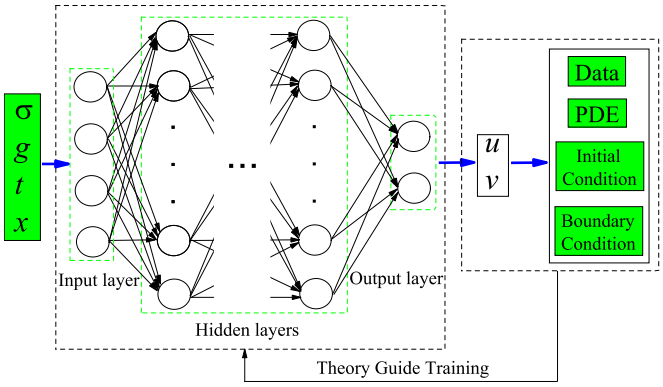

FIG. 3. Illustration of the TgNN, which contains an input layer, hidden layers, and an output layer. Each circle represents a neuron, and each arrow connecting them represents a functional mapping between the two neurons. The input includes σ, g, t, and x. The output contains u and v, which are the real and imaginary parts of the wave function, respectively. The deep learning process is not only driven by data, but more importantly by its Gross-Pitaevskii equation (PDE), initial condition and boundary condition.

and

$$R^v_{data}=\frac{1}{N_d}\sum_{k=1}^{N_d}\left[v_k^{NN}(g,\sigma,x,t;\mathbf{W},\mathbf{b})-v_k(g,\sigma,x,t;\mathbf{W},\mathbf{b})\right]^2, \tag{9}$$

where u_k^{NN} and u_k are the prediction of the neural network and the true data, respectively. Obviously, the smaller the R^u_{data} is, the closer the predicted result of u is to the true data. For R^v_{data}, it is the same. Similarly, the smaller the R^Γ_f is, the more the result predicted by the neural network enables Eq. (5) to be established. The smaller both R^Γ_I and R^Γ_B are, the closer the initial and boundary results predicted by the neural network are to the true initial and boundary conditions, respectively. Finally, the loss function of TgNN can be represented by:

$$\begin{aligned}L_{\mathrm{TgNN}} = {}& \lambda_{data}\left(R^u_{data}+R^v_{data}\right)+\lambda_I\left(R^u_I+R^v_I\right)\\ &+\lambda_B\left(R^u_B+R^v_B\right)+\lambda_f\left(R^u_f+R^v_f\right),\end{aligned} \tag{10}$$

where λ_{data}, λ_f, λ_I, and λ_B are the weights of each residual term, which can be tuned during the training process, so as to achieve more desirable performance based on the value of the loss function.

The structure of the TgNN is described in detail in Fig. 3. For comparison, we here set four hidden layers with 240 neurons in each layer, and choose softplus $f(\delta)=\ln(1+e^\delta)$ as the activation function both in TgNN and DNN. We randomly extracted $x_i\in(-40,40)$, $i=1,\ldots,320$ and $t_i\in(0,40)$, $i=1,\ldots,100$ in the real-time evolution process of each dynamic solution marked by black dots in Fig. 2 as the training data set. For the TgNN, 6000 and 10000 collocation points are used to impose boundary and initial condition constraints, respectively. 10000 collocation regions are sampled from the whole domain (g,σ,x,t) to impose the Eq. (5) constraint, with their center coordinates determined using the Latin hypercube sampling strategy.

IV. RESULTS

In order to examine the performance of TgNN, we plot Fig. 4, where we have set $\sigma=20$ and g as increasing. The first three rows represent the true results, predictions of DNN, and predictions of TgNN, respectively; the colors represent atomic density $|\psi|=(u^2+v^2)^{1/2}$ in all of them. The last two rows represent the relative errors between the true results and predictions from DNN and TgNN, respectively; their strengths are denoted by the colors. Furthermore, because our focus here is the first-order RWs of this system, in order to more clearly demonstrate the superiority of the TgNN model, we plot Fig. 5 to perform a more detailed investigation of their structures, which are shown in Fig. 4. The first two rows represent the predictions of the spatial structure at the peak of the first-order RWs from DNN and TgNN, respectively, and they correspond to the first white dotted lines in Figs. 4(a)–4(i). The last two rows represent the predictions of the spatial structure at the peak of weak-first-order RWs from DNN and TgNN, respectively, and they correspond to the second white dotted lines in Figs. 4(a)–4(i). The solid blue lines represent the true data in all cases.

From the first column of Fig. 4, it can be seen that if $g=-5.5$ and $\sigma=20$, in which the data are known and have been used for network training (see Fig. 2), both DNN and TgNN, as expected, can accurately give the results of dynamics, including the first-order RWs, with almost no errors. Accordingly, the first-order RWs and weak-first-order RWs are accurately predicted by both DNN and TgNN, as shown in the first column of Fig. 5. If $g=-5.4$ increased slightly, in which the data are unknown and have not been used for network training, both DNN and TgNN can still predict their dynamics, but the results from DNN are slightly inaccurate, as shown in the second column of Fig. 4. Thus, the first-order RWs predicted by TgNN are accurate [see Fig. 5(e)], while the others are slightly inaccurate. However, if $g=-5$ increased more, the result error of DNN is considerable, and the dynamic structure is destroyed, as shown by the second white dotted line in Fig. 4(f). Consequently, the results of DNN become unreliable [see Fig. 5(c) and 5(i)]; whereas, the result of TgNN remains relatively accurate, especially the first-order RW about which are concerned [see Fig. 5(f)]. These results demonstrate that the TgNN approach is superior to the traditional DNN due to the theory guidance of such problem.

In order to determine whether the TgNN approach can assist us to discover the first-order RWs that have not been discovered under other parameters in such system, we next randomly select three parameter points $(g,\sigma)=(-3.4,23)$, $(-5,26)$, and $(-1.8,31)$ marked by red stars within the valid parameter plane of Fig. 2 for examination, and plot Figs. 6–7. It is worth noting that if we do not use the TgNN for prediction, it is unknown whether the RWs can be excited, and what their structures will be, until the relevant calculations are performed by traditional RTEM under these parameters. In Fig. 6, the first three rows represent the true results, predictions of DNN, and predictions of TgNN, respectively; the last two rows represent the relative errors between the true results and predictions from DNN and TgNN, respectively. The predictions of first-order RWs from DNN

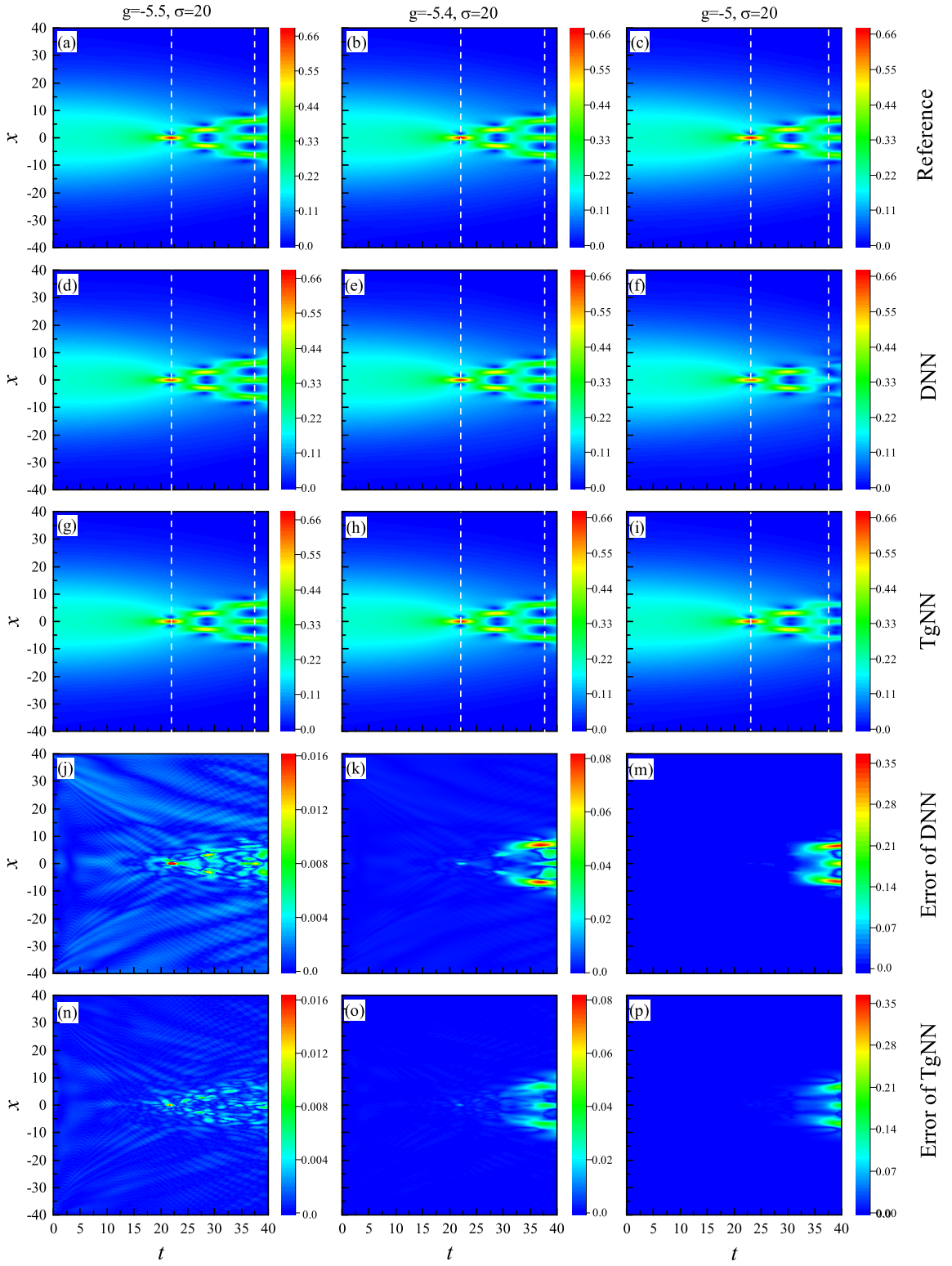

FIG. 4. Comparison of predicted results of dynamic properties between the DNN and the TgNN. The first three rows represent the true results, predictions of DNN, and predictions of TgNN, respectively. The colors represent atomic density $|\psi| = (u^2 + v^2)^{1/2}$, and the white lines correspond to the results of Fig. 5. The last two rows represent the relative errors between the true results and predictions from DNN and TgNN, respectively. In all cases, the parameters (g, σ) correspond to the red stars in Fig. 2.

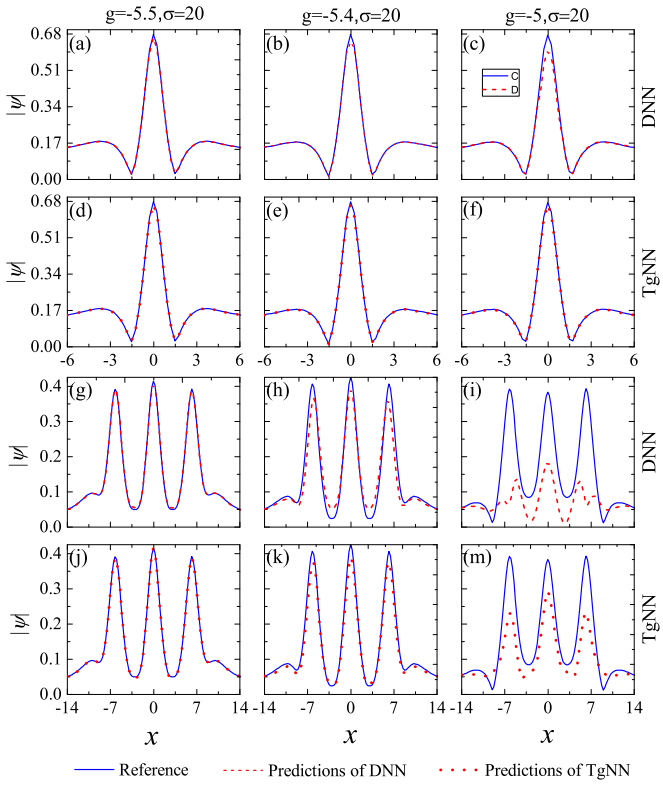

FIG. 5. The first two rows represent the predictions of first-order RWs from DNN and TgNN, respectively. They correspond to the first white dotted lines in Figs. 1(a)–1(i). The last two rows represent the predictions of weak-first-order RWs from DNN and TgNN, respectively. They correspond to the second white dotted lines in Figs. 1(a)–1(i). In all cases, the solid blue lines represent the true data, and the parameters (α, γ) correspond to Fig. 4.

and TgNN are presented in Fig. 7, and they correspond to the white dotted lines in Fig. 6. The solid lines indicate the true data in all cases.

It can been seen from Fig. 6, if $(g, \sigma) = (-3.4, 23)$ and $(-5, 26)$, the first-order RWs can be excited, which are able to be predicted by both DNN and TgNN, as shown in the first two columns. However, the prediction results of TgNN for the first-order RW structure are significantly better than those of DNN, even for the point $(-5, 26)$ far away from the training parameter point. This is demonstrated in more detail in Fig. 7. In contrast, if $(g, \sigma) = (-1.8, 31)$, no RWs can be found within our observation period, and as anticipated, the performance of TgNN is also better than that of DNN, as shown in the third column of Fig. 6. In other words, not all parameter conditions can lead to the generation of RWs in this parameter plane, and they could be excited only in special region. Therefore, based on the TgNN, the parameter plane is further clearly divided by a dotted line into two regions: phase I and phase II. In phase I, the RWs can be excited, but they cannot be excited in phase II. This is in close accordance with the results obtained by the traditional real-time iterative calculation method represented by the blue upper triangles. These results show that the TgNN can be successfully used, and replace traditional methods, to explore the new physics of first-order RWs in such system.

It is worth noting that, compared with the TgNN approach, it will take approximately 48 s to obtain the structures of the RWs above using the RTEM in this system; whereas, the TgNN approach requires only approximately 0.85 s. This greatly reduces the time cost of exploring the RW structures of such system. Furthermore, if one intends to explore the properties of RWs outside of the plane of the current valid parameter, transfer learning [53] may constitute an appropriate choice, which may need less time to retrain a new TgNN with new real-time evolution data outside of this plane based on the above results.

Finally, in order to quantitatively evaluate these results, we introduce two metrics. They are relative L_2 error and coefficient of determination (R^2 score), which are defined as follows:

$$L_2 = \frac{\| \varepsilon^{NN} - \varepsilon \|_2}{\| \varepsilon \|_2}, \quad R^2 = 1 - \frac{\Sigma_{k=1}^{N_R} (\varepsilon_k^{NN} - \varepsilon_k)^2}{\Sigma_{k=1}^{N_R} (\varepsilon_k - \bar{\varepsilon})^2}, \tag{11}$$

where $\varepsilon = |\psi| = (u^2 + v^2)^{1/2}$ and $\varepsilon^{NN} = |\psi^{NN}| = [(u^{NN})^2 + (v^{NN})^2]^{1/2}$; $\| \cdot \|_2$ denotes the standard Euclidean norm; N_R is the total number of evaluation points; and $\bar{\varepsilon}$ is the average value of ε. Table I is calculated based on these two metrics. It can be seen from Table I that if the data of parameter point has been used for neural network training, there is only a negligible difference in L_2 and R^2 between DNN and TgNN, such as parameter point $(g = -5.5, \sigma = 20)$. In contrast, if the parameter points are far away from the training data points, L_2 of TgNN is always smaller than that of DNN, while R^2 is the opposite of this. These results once again quantitatively demonstrate the superior performance of TgNN and its reliability to explore the RWs of a physical system.

TABLE I. Comparison of L_2 and R^2 for DNN and TgNN.

	L_2		R^2	
(g, σ)	TgNN	DNN	TgNN	DNN
$(-5.5, 20)$	7.3119e-03	1.3655e-02	9.9989e-01	9.9964e-01
$(-5.4, 20)$	3.1929e-02	5.5979e-02	9.9806e-01	9.9404e-01
$(-5, 20)$	1.1989e-01	2.1601e-01	9.7246e-01	9.1061e-01
$(-3.4, 23)$	1.0495e-02	1.4551e-02	9.9974e-01	9.9950e-01
$(-5, 26)$	1.9430e-02	2.4744e-02	9.9900e-01	9.9837e-01
$(-1.8, 31)$	2.0448e-02	2.3277e-02	9.9844e-01	9.9798e-01

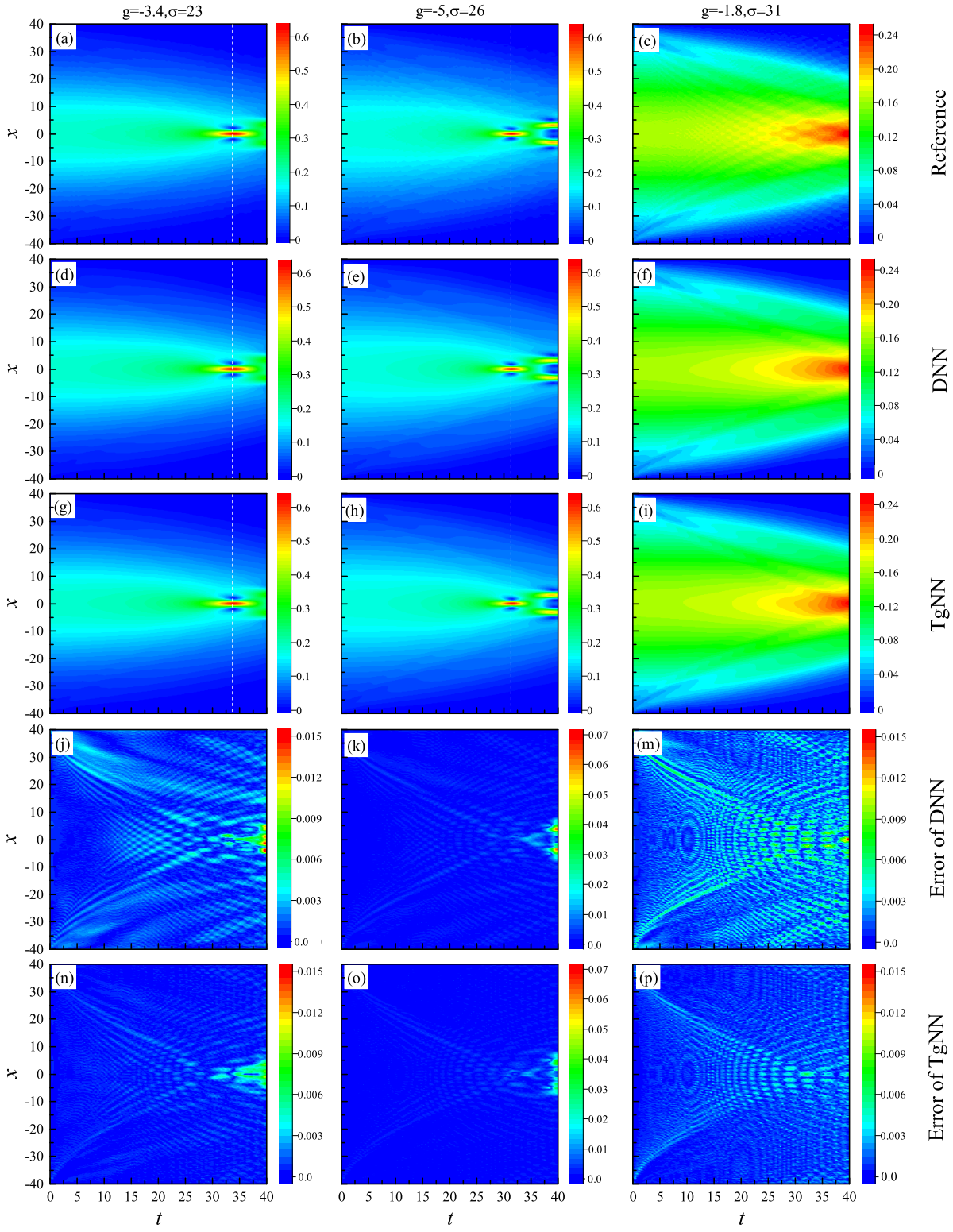

FIG. 6. Comparison of predicted results of dynamic properties between the DNN and the TgNN. The first three rows represent the true results, predictions of DNN, and predictions of TgNN, respectively. The colors represent atomic density $|\psi| = (u^2 + v^2)^{1/2}$, and the white lines correspond to the results of Fig. 7. The last two rows represent the relative errors between the true results and predictions from DNN and TgNN, respectively. In all cases, the parameters (g, σ) correspond to the red stars in Fig. 1.

V. CONCLUSION

By incorporating scientific knowledge, such as GP equations, boundaries and initial conditions, into the neural network, a theory-guided neural network (TgNN) model is constructed in this work to explore the rogue waves (RWs) of one-dimensional Bose-Einstein condensates. Through the TgNN, deep learning is not only driven by data, but more

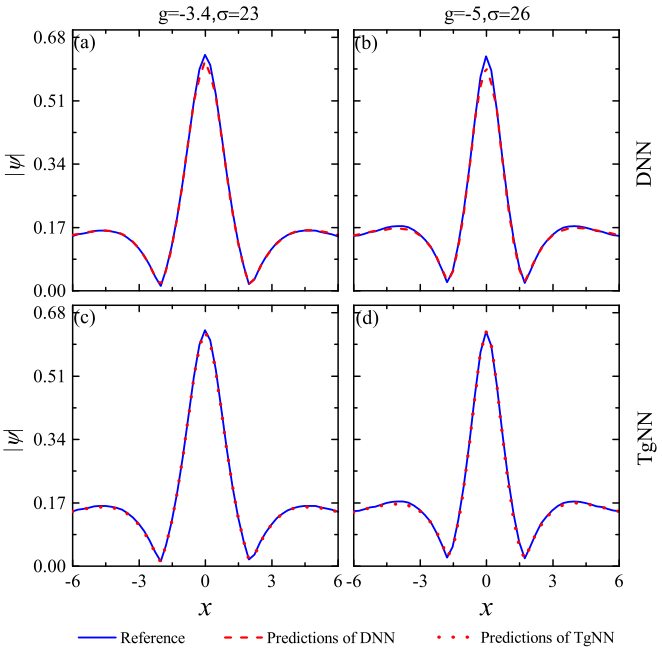

FIG. 7. The two rows represent the predictions of first-order RWs from DNN and TgNN, respectively. They correspond to the white dotted lines in Figs. 5(a)–5(b) and 5(c)–5(d). In all cases, the solid blue lines represent the true solutions, and the parameters (α, γ) correspond to Fig. 6.

importantly by physical laws, which has enabled the model to achieve better accuracy, generalization ability, and robustness. The results show that such method is superior to the traditional deep neural network (DNN) because of theory guidance of underlying problems, and it is able to directly give any unknown excited location, timing, and structure of the first-order RWs using only a small amount of dynamic evolution data gained from real-time evolution methods (RTEM) as the training data, without an onerous and repetitive tedious step-by-step iterative calculation process. Furthermore, based on such method, a phase transition boundary is also discovered, which clearly distinguishes the first-order RW phase from the non-RW phase. This means that based on the tedious RTEM method the TgNN model can practically become a faster and more efficient universal method to explore unusual phenomena of RW in a system. Finally, the TgNN approach not only greatly conserves computing resources and time, but also constitutes a promising and universal technique for discovering novel structures of RWs in other parameterized nonlinear systems.

Although we focus here on the universal BECs, our results may also be applicable in spin-orbit coupling BECs, multicomponent BECs, and two-dimensional or even three-dimensional BECs. The more complex the system being explored, the more difficult and time consuming it is to explore rogue waves with traditional methods, thus the TgNN method will have more advantages.

ACKNOWLEDGMENT

This work is supported by special funds from the Peng Cheng Laboratory.

[1] C. Kharif, E. Pelinovsky, and A. Slunyaev, *Rogue Waves in the Ocean* (Springer, Heidelberg, 2009).
[2] D. H. Peregrine, J. Aust. Math. Soc. Ser. B. Appl. Math. **25**, 16 (1983).
[3] P. Müller, Ch. Garrett, and A. Osborne, Oceanogr. **18**, 66 (2005).
[4] L. Stenflo and P. K. Shukla, J. Plasma Phys. **75**, 841 (2009).
[5] M. Erkintalo, G. Genty, and J. M. Dudley, Opt. Lett. **34**, 2468 (2009).
[6] W. M. Moslem, P. K. Shukla, and B. Eliasson, Europhys. Lett. **96**, 25002 (2011).
[7] A. N. Ganshin, V. B. Efimov, G. V. Kolmakov, L. P. Mezhov-Deglin, and P. V. E. McClintock, Phys. Rev. Lett. **101**, 065303 (2008).
[8] Y. V. Bludov, V. V. Konotop, and N. Akhmediev, Phys. Rev. A **80**, 033610 (2009).
[9] M. Shats, H. Punzmann, and H. Xia, Phys. Rev. Lett. **104**, 104503 (2010).
[10] N. Akhmediev, A. Ankiewicz, and M. Taki, Phys. Lett. A **373**, 675 (2009).
[11] N. Akhmediev, A. Ankiewicz, and J. M. Soto-Crespo, Phys. Rev. E **80**, 026601 (2009).
[12] L. C. Zhao, L. Duan, P. Gao, and Z. Y. Yang, Europhys. Lett. **125**, 40003 (2019).
[13] A. Chabchoub, N. P. Hoffmann, and N. Akhmediev, Phys. Rev. Lett. **106**, 204502 (2011).
[14] H. Bailung, S. K. Sharma, and Y. Nakamura, Phys. Rev. Lett. **107**, 255005 (2011).
[15] M. J. Ablowitz and J. T. Cole, Phys. Rev. Lett. **127**, 104101 (2021).
[16] A. Romero-Ros, G. C. Katsimiga, S. I. Mistakidis, B. Prinari, G. Biondini, P. Schmelcher, and P. G. Kevrekidis, Phys. Rev. A **105**, 053306 (2022).
[17] A. Chabchoub, N. Hoffmann, M. Onorato, and N. Akhmediev, Phys. Rev. X **2**, 011015 (2012).
[18] B. Frisquet, B. Kibler, and G. Millot, Phys. Rev. X **3**, 041032 (2013).
[19] F. Sgrignuoli, Y. Chen, S. Gorsky, W. A. Britton, and L. Dal Negro, Phys. Rev. B **103**, 195403 (2021).
[20] A. Chowdury and W. Chang, Phys. Rev. Res. **3**, L032060 (2021).
[21] F. Meng, C. Lapre, C. Billet, T. Sylvestre, J. M. Merolla, C. Finot, and J. M. Dudley, Nature Commun. **12**, 1 (2021).
[22] G. Marcucci, D. Pierangeli, and C. Conti, Phys. Rev. Lett. **125**, 093901 (2020).
[23] S. S. Gopalakrishnan, K. Panajotov, M. Taki, and M. Tlidi, Phys. Rev. Lett. **126**, 153902 (2021).
[24] K. Talouneh, R. Kheradmand, G. Tissoni, and M. Eslami, Phys. Rev. A **105**, 013501 (2022).
[25] Z. Y. Sun and X. Yu, Phys. Rev. E **103**, 062203 (2021).
[26] J. Chen and D. E. Pelinovsky, Phys. Rev. E **103**, 062206 (2021).
[27] Z. Zhou and Z. Yan, Commun. Theor. Phys. **73**, 105006 (2021).

[28] L. Wang and Z. Yan, Physica D **428**, 133037 (2021).
[29] L. Wang and Z. Yan, Phys. Lett. A **404**, 127408 (2021).
[30] M. I. Jordan and T. M. Mitchell, Science **349**, 255 (2015).
[31] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature (London) **521**, 436 (2015).
[32] X.-D. Bai, Jie Zhao, Y.-Y. Han, J.-C. Zhao, and J.-G. Wang, Phys. Rev. B **103**, 134203 (2021).
[33] J. Carrasquilla and R. G. Melko, Nature Phys. **13**, 431 (2017).
[34] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
[35] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B **96**, 195145 (2017).
[36] X.-Y. Dong, F. Pollmann, and X.-F. Zhang, Phys. Rev. B **99**, 121104(R) (2019).
[37] E. P. Van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nature Phys. **13**, 435 (2017).
[38] P. Huembeli, A. Dauphin, and P. Wittek, Phys. Rev. B **97**, 134109 (2018).
[39] J. F. Rodriguez-Nieva and M. S. Scheurer, Nature Phys. **15**, 790 (2019).
[40] N. L. Holanda and M. A. R. Griffith, Phys. Rev. B **102**, 054107 (2020).
[41] Y. Zhang, P. Ginsparg, and E.-A. Kim, Phys. Rev. Res. **2**, 023283 (2020).
[42] Y. Long, J. Ren, and H. Chen, Phys. Rev. Lett. **124**, 185501 (2020).
[43] M. S. Scheurer and R.-J. Slager, Phys. Rev. Lett. **124**, 226401 (2020).
[44] A. Dawid, J. Arnold, B. Requena *et al.*, arXiv:2204.04198.
[45] R. Xu, D. Zhang, M. Rong, and N. Wang, J. Comput. Phys. **436**, 110318 (2021).
[46] M. Germain, H. Pham, and X. Warin, arXiv:2101.08068.
[47] M. Emadi, R. Taghizadeh-Mehrjardi, A. Cherati, M. Danesh, A. Mosavi, and T. Scholten, Remote Sens. **12**, 2234 (2020).
[48] X. D. Bai and D. Zhang, Phys. Rev. A **104**, 063316 (2021).
[49] Y. Khoo, J.-F. Lu, and L.-X. Ying, Eur. J. Appl. Math. **32**, 421 (2021).
[50] M. Raissi and G. E. Karniadakis, J. Comput. Phys. **357**, 125 (2018).
[51] M. Raissi, P. Perdikaris, and G. E. Karniadakis, J. Comput. Phys. **378**, 686 (2019).
[52] J.-C. Pu and Y. Chen, Chaos, Solitons Fractals **160**, 112182 (2022).
[53] N. Wang, D. Zhang, H. Chang, and H. Li, J. Hydrol. **584**, 124700 (2020).
[54] D. M. Gangardt and G. V. Shlyapnikov, Phys. Rev. Lett. **90**, 010401 (2003).
[55] J. E. Williams, The preparation of topological modes in a strongly-coupled two-component Bose-Einstein condensate, PhD. thesis, University of Colorado, 1999.
[56] F. Dalfovo and G. V. Shlyapnikov, Rev. Mod. Phys. **71**, 463 (1999).
[57] A. Tikan, C. Billet, G. El, A. Tovbis, M. Bertola, T. Sylvestre, F. Gustave, S. Randoux, G. Genty, P. Suret, and J. M. Dudley, Phys. Rev. Lett. **119**, 033901 (2017).
[58] E. G. Charalampidis, J. Cuevas-Maraver, D. J. Frantzeskakis, and P. G. Kevrekidis, Rom. Rep. Phys. **70**, 504 (2018).
[59] W. Bhimji, S. A. Farrell, T. Kurth, M. Paganini, Prabhat, and E. Racah, J. Phys.: Conf. Ser. **1085**, 042034 (2018).