


**Random sampling neural network for quantum many-body problems**Chen-Yu Liu  and Daw-Wei Wang<sup>1,2,3</sup><sup>1</sup>*National Center for Theoretical Sciences, Hsinchu 30013, Taiwan*<sup>2</sup>*Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan*<sup>3</sup>*Center for Quantum Technology, National Tsing Hua University, Hsinchu 30013, Taiwan* (Received 21 November 2020; revised 2 March 2021; accepted 20 April 2021; published 4 May 2021)

The eigenvalue problem of quantum many-body systems is a fundamental and challenging subject in condensed-matter physics because the dimension of the Hilbert space (and hence the required computational memory and time) grows exponentially as the system size increases. A few numerical methods have been developed for some specific systems but may not be applicable in others. Here we propose a general numerical method, random sampling neural networks (RSNNs), to utilize the pattern recognition technique for the random sampling matrix elements of an interacting many-body system via a self-supervised learning approach. Several exactly solvable one-dimensional models, including the Ising model with a transverse field, the Fermi-Hubbard model, and the spin- $\frac{1}{2}$  XXZ model, are used to test the applicability of RSNN. Pretty high accuracy ( $>96\%$ ) of energy spectra, magnetization, critical exponents, etc. can be obtained within the strongly correlated regime or near the quantum phase-transition point, even the corresponding RSNN models are trained in the weakly interacting regime. After including data augmentation and transfer learning methods, RSNN can be further applied to systems of larger sizes with much less training data if only pretrained in a smaller system. Our results demonstrate the possibility to combine the existing numerical methods and the RSNN to explore quantum many-body problems in a much wider parameter space.

DOI: [10.1103/PhysRevB.103.205107](https://doi.org/10.1103/PhysRevB.103.205107)**I. INTRODUCTION**

It has been a long-standing challenge in condensed-matter physics that the eigenvalue problem of a many-body system is in general not accessible because of the exponentially huge Hilbert space of the associated quantum many-body Hamiltonian. Only very few simple models in low-dimensional space can be solved exactly due to their higher order symmetries [1–3]. As a result, various analytical and numerical methods are developed for certain specific systems, including perturbation theory [4,5], renormalization group [6–8], bosonization [9,10], quantum Monte Carlo [11–13], density matrix renormalization group [14,15], tensor networks [16,17], etc. Within these analytic or numerical methods, there are also many exquisite techniques developed for solving some specific many-body problems in some parameter regimes. Furthermore, in the recent rapidly growing development of machine-learning approaches [18], certain unsupervised learning methods, such as neural network quantum state (NQS), are found to have better results than ordinary variational methods in the calculation of the ground-state and excited-state energies [19,20] through its undetermined parameters of a restricted Boltzmann machine. However, in these various approaches, each data point is calculated independently according to the associate system parameters, and therefore it may cost a lot of computational resources to get a complete phase diagram.

From the data-driven machine-learning point of view, on the other hand, this problem could be investigated from differ-

ent perspectives. Instead of unsupervised learning approaches, which are similar to variational methods, one could also train a model based on existing results in a well-known parameter regime and then apply this model to other regimes. Most applications along this line are the classification and/or feature extraction of various phase transitions through the supervised learning approach [21–24]. The basic concept is to train a model to learn the identities (i.e., labels) of different phases in some parameter regimes and apply it to define the phase boundary in the middle regime. Such a numerical approach is possible because a horizontal relationship between features in different parameter regimes could be learned (more precisely, fit by a complicated function through machine learning) during the training process. However, since the input features are usually from experimental data or other physical quantities, and the labels of these phases are in fact artificial labeling for the purpose of classification [say,  $=(0, 1)$  for phase A and  $=(1, 0)$  for phase B], such a pattern-recognition scenario could not provide sufficient information for the understanding of a many-body system. After all, the “nature” of these phases should be associated with the relationship between these physical quantities, while such a relationship is in general too complicated to be calculated in most systems.

Combining the advantages of the two approaches mentioned above, in this paper, we propose a self-supervised machine-learning method, random sampling neural network (RSNN), to study a general many-body problem. Motivated by pattern recognition and self-supervised methods in computer vision [25–27], we treat the many-body Hamiltonian

as a huge two-dimensional (2D) “system image.” Random sampling matrix elements (i.e., “patches”) are collected from this system image as the input features to train a convolutional neural network (CNN) in the training regime. The labels of these features are physical quantities obtained by the system Hamiltonian at the same parameter, reflecting the spirit of a self-supervised learning process. We show that the accuracy of such a simulation in the test regime can be systematically improved by increasing the amounts of training data, while the computation time just scales linearly to the system size, independent of the physical quantities to investigate. We use several one-dimensional (1D) exactly solvable models, including an Ising model with a transverse field, a Fermi-Hubbard model, and a spin- $\frac{1}{2}$  XXZ model, to demonstrate the applicability of RSNN in the strongly correlated regime, if only the model is properly trained by known results (obtained by other numerical methods) in the weakly interacting regime. Furthermore, when the RSNN model is pretrained by data in a smaller system, we could use transfer learning method to obtain results of a much larger system with much less training data. Our results therefore demonstrate that RSNN is an efficient data-driven and complementary approach to the existing numerical methods for the study of quantum many-body problems.

In the rest of this paper, we first introduce the basic concept and hypothesis of RSNN in Sec. II, and then use the 1D Ising Model with a transverse field as the first example of a RSNN in Sec. III. In Sec. IV, we then use 1D IMTF to systematically investigate how the accuracy and computation time change for different hyperparameters of RSNN. Similar results should be also expected for other physics models. We then apply RSNN to predict the whole energy spectrum of a 1D Fermi-Hubbard Model in the strongly correlated regime in Sec. V. In Sec. VI, we further apply RSNN to predict the quantum phase-transition point of the 1D XXZ model and investigate its quantum critical exponent. After applying RSNN in different models, in Sec. VII, we utilize the data augmentation and transfer learning method to improve RSNN results for fewer training data and/or for larger system sizes. We then summarize our results in Sec. VIII, and provide a Github code. Further details of the RSNN models are shown in the Appendix.

## II. BASIC CONCEPT AND HYPOTHESIS

The concept of RSNN is motivated by machine-learning methods developed for pattern recognition: The system Hamiltonian  $\hat{\mathcal{H}}$  can be treated as a 2D system image, after represented by a Hermitian matrix  $\mathbf{H}$  with matrix elements  $H_{\mu\nu} = \langle \phi_\mu | \hat{\mathcal{H}} | \phi_\nu \rangle$ . Here  $\{|\phi_\nu\rangle\}$  is a complete and orthonormal basis. Each  $H_{\mu\nu}$  is like a “two-color pixel” with its real and imaginary parts. As a result, the process to calculate any physical quantities (for example, its eigenenergies  $E_n$  or any expectation value of the ground state) is thus equivalent to deriving the functional relationship between the matrix elements and these quantities, i.e.,  $F(\mathbf{H}[\lambda]) = F(\{H_{\mu\nu}(\lambda)\}) = \{E_n(\lambda)\}$ , where  $\lambda$  stands for a system parameter in control (for example, external magnetic field or coupling strength). From the machine-learning point of view, interestingly, solving such a functional relationship  $F$  is thus equivalent to training a neural

network model ( $F_{\text{NN}}$ ) to simulate this complicated function, i.e.,

$$F_{\text{NN}}(\mathbf{H}[\lambda]) \Rightarrow F(\mathbf{H}[\lambda]) = \{E_n(\lambda)\}. \quad (1)$$

According to the universal approximation theorem [28], the difference between the approximate function  $F_{\text{NN}}$  and the true function  $F$  can be infinitesimal if the number of artificial neurons (and hence the fitting parameters) and training data used for  $F_{\text{NN}}$  are large enough.

In a general many-body system, however, the dimension of such a system image (i.e., the matrix representation of the system Hamiltonian) grows exponentially as the system size increases and therefore cannot be simulated efficiently. To overcome this problem, here we propose a random sampling method, where a large image could still be recognized if only patches of this image are used during the training process. More precisely, we first randomly select  $M$  basis vectors from the full many-body basis to construct an  $M \times M$  sampling matrix  $\mathbf{H}_S^{(m)}$  and repeat this sampling  $N_S$  times ( $m = 1, \dots, N_S$ ). The selected basis for each time can be different from each other. These random sampling matrices therefore form a collection of patches of the original system image and hence contain partial information of the full many-body system.

We could then combine these random sampling data and the concepts of a self-supervised learning model, which is trained by the internal properties of a system rather than external labels [29], and propose the following *random sampling hypothesis*: The patches of the full system image extracted from random sampling basis can be used as the input features of a neural network model so that, in a given training regime ( $\lambda \in R_{\text{train}}$ ), the obtained random sampling function  $F_{\text{RSNN}}$  can simulate the target physical quantities via a self-supervised learning process within a small deviation:

$$|F_{\text{RSNN}}(\{\mathbf{H}_S^{(m)}[\lambda], \mathbf{b}_S^{(m)}[\lambda]\}) - \{E_n(\lambda)\}| < \epsilon. \quad (2)$$

Here the upper bound of their difference,  $\epsilon$ , can be reduced only if the amounts of artificial neurons and/or the training data are increased. Its application in the test regime ( $\lambda \in R_{\text{test}}$ ) can therefore also provide reliable estimates if  $R_{\text{test}}$  is not too far from  $R_{\text{train}}$ .

Before applying this hypothesis to a realistic physical problem, we have to emphasize that the collection of sampling basis ( $\{\mathbf{b}_S^{(m)}[\lambda]\}$ ) can be different for each time, so that the neural network could be enforced to simulate the eigenvalue-solving problem based on the partial information of the selected basis. A similar scenario can be also applied to the calculation of other physical quantities, such as magnetization or a spectral function. In the rest of this paper, we provide numerical calculations to support this hypothesis and investigate its application in different many-body problems.

## III. GROUND STATE AND MAGNETIZATION OF ONE-DIMENSIONAL ISING MODEL WITH A TRANSVERSE FIELD

To demonstrate the applicability of RSNN, we first take a 1D spin- $\frac{1}{2}$  Ising model with a transverse field (IMTF) [1] as

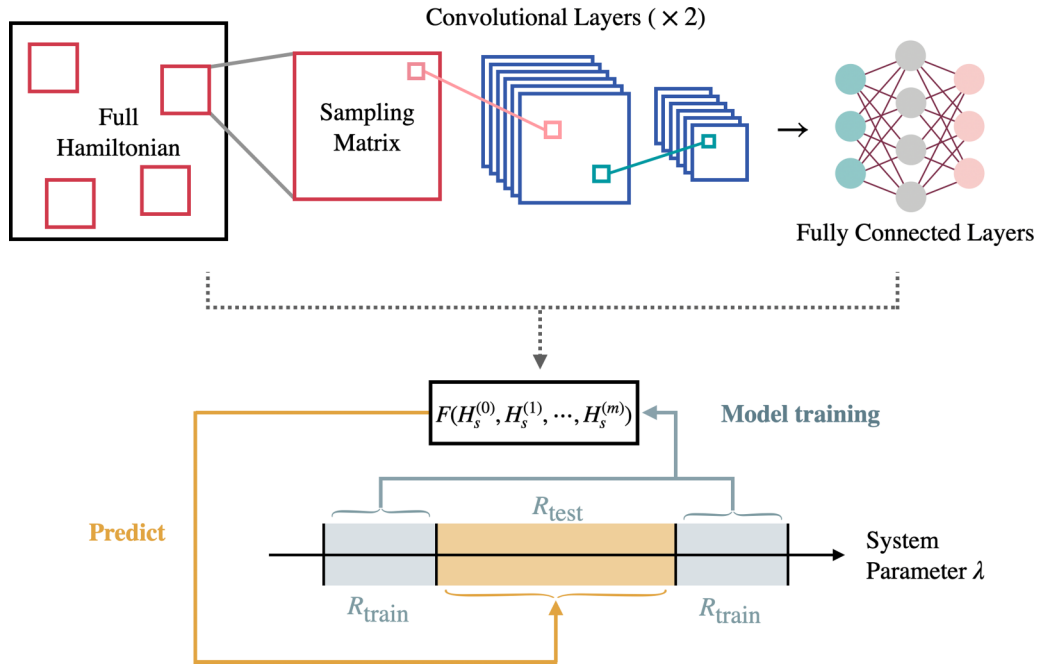


FIG. 1. A typical flowchart of a RSNN model. The input data are  $M \times M$  matrices  $\mathbf{H}_s^{(m)}$ , randomly sampled from the original full system Hamiltonian (see text). After a standard self-supervised learning by a CNN model with known results in the training regime ( $R_{\text{train}}$ ), the obtained RSNN model can be used to predict physical results in the test regime ( $R_{\text{test}}$ ). In most physical problems, the training and test regimes can be defined by a continuous parameter  $\lambda$ .

an example for systematic studies. The system Hamiltonian is known to be

$$\hat{\mathcal{H}}_{\text{IMTF}} = -J \sum_i^N \sigma_i^z \hat{\sigma}_{i+1}^z - h \sum_i^N \hat{\sigma}_i^x, \quad (3)$$

where  $\hat{\sigma}_{x,y,z}$  are Pauli matrices,  $J > 0$  is the spin coupling between the nearest-neighbor site,  $h$  is the transverse field strength, and  $N$  is the total number of spins. 1D IMTF with the periodic boundary condition is exactly solvable though the Jordan-Wigner transformation [30], and therefore could be a good example to test the application of our RSNN. In the thermodynamic limit, the ground state is a doubly degenerate ferromagnetic phase for  $h < J$  and becomes a nondegenerate paramagnetic phase for  $h > J$ . We could then define  $\lambda \equiv h/(h + J) \in (0, 1)$  as a dimensionless system parameter (see Fig. 1) to measure such a phase transition.

In Fig. 2(a) we show the results of RSNN for the three lowest eigenenergies of  $N = 12$ . Note that the energy is scaled by  $J + h$  in order to have a better expression of the energy in the both sides of training regime. We select  $N_{\text{train}} = 10000$  values of  $\lambda$  from the training regime (colored background) and generate  $N_s = 200$  sampling matrices (with a dimension  $M = 10$ ) for each  $\lambda$  as the training data, labeled by the exact eigenstate energies. Since the ground state is pretty well known in the regime as  $\lambda \rightarrow 0$  and  $\lambda \rightarrow 1$ , we choose the training regime on both sides and predict the results in the middle (test) regime,  $R_{\text{test}} = (0.3, 0.7)$ , where a first-order quantum phase transition is expected to appear around  $\lambda = 0.5$  in the thermodynamic limit. Comparing the results of RSNN with the exact results in the test regime, we find the accuracy  $\text{Acc} \equiv 1 - \text{Ave}[\sum_{i=0}^2 |E_i^{\text{RSNN}} -$

$E_i^{\text{ED}}|/|E_i^{\text{ED}}|] = 99.43\% \pm 0.19\%$ . Here  $E_i^{\text{RSNN}}$  and  $E_i^{\text{ED}}$  are the  $i$ th eigenvalues obtained by RSNN and the exact solution, respectively, and  $\text{Ave}[\dots]$  is the average taken over the whole test regime for five independent calculations. The inset of Fig. 2(a) shows the predicted magnetization by RSNN for  $N = 30$ , where the training values of magnetization are calculated by matrix product state (MPS) method [31,32]. Notice that the eigenstate energies and magnetization predicted by RSNN are very close to the exact results. Below we use this model to investigate the accuracy and efficiency of RSNN in various conditions. Details of these model parameters are shown in Appendix A.

In Fig. 2(b), we show the average computation time  $t_{\text{RSNN}}$  of each data for RSNN as a function of system size  $N$ . The computation time  $t_{\text{MPS}}$  by MPS is also shown together for comparison. Here  $t_{\text{RSNN}}$  is calculated by adding all the generation times of random sampling matrices (for both the training data and the test data) as well as the training time of the RSNN model, and then dividing by the total number of test data ( $N_{\text{test}} = 100$ ). For comparison, we also show the calculation time of MPS (dashed line) in the same plot. We find that, as opposed to the exponentially growing time by exact diagonalization (not shown here) or the long computation time of MPS in the large- $N$  limit,  $t_{\text{RSNN}}$  grows much more slowly and linearly in the large- $N$  limit, while the accuracy of output is still above 99% for the eigenvalues (not shown here) and above 92% for the magnetization even for  $N > 100$ . Note that, since the required computation time for the feature generation and model training of RSNN is almost the same no matter how many test data are generated, RSNN could be a very efficient method to generate many more test data in the whole parameter regime

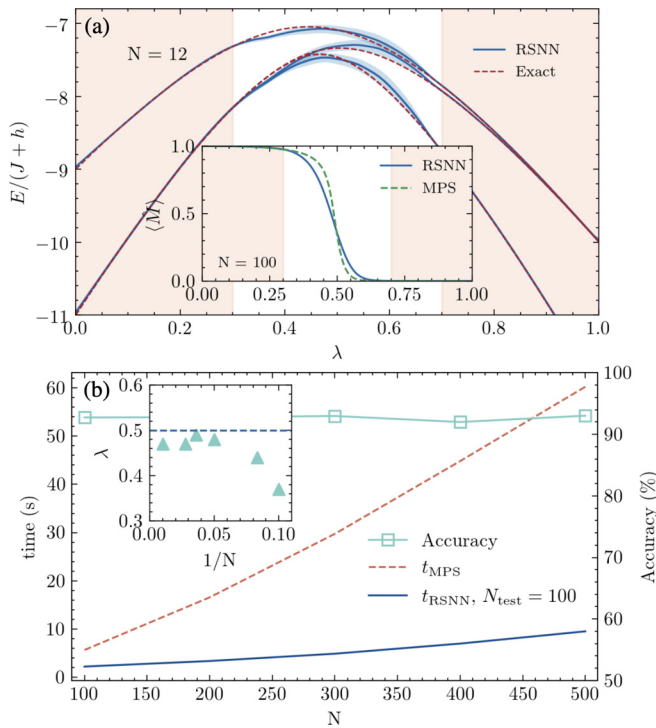


FIG. 2. (a) Prediction result of the lowest three eigenenergies of 1D IMTF by RSNN (solid lines) for  $N = 12$ . The training regime,  $R_{\text{train}} = (0, 0.3) \cup (0.7, 1.0)$ , is shown by the colored background, and the test regime,  $R_{\text{test}} = (0.3, 0.7)$  is shown by the white background. The blue-gray area near the RSNN result indicates the uncertainty, resulting from five independent calculations. The results obtained by the exact diagonalization (dashed lines) are shown together for comparison. The inset shows the predicted magnetization (solid line) for  $N = 30$  compared with the results obtained by MPS (dashed line). Here we use  $N_{\text{train}} = 10\,000$ ,  $N_S = 200$ , and  $M = 10$  for the training process (see text). Note that we have scaled the energy in unit of  $J + h$  in order to provide a more balanced expression of eigenstate energies in the training regime. Panel (b) shows the average computation time (solid line) to generate a test data by RSNN for  $N_{\text{test}} = 100$  data points (see text). The computation time by MPS (dashed lines) as a function of system sizes are also shown together for comparison. The accuracy of the magnetization of the RSNN method is also shown together by open squares. The inset shows the finite-size scaling of the phase-transition point, which is defined when the separation between the lowest two eigenenergies is larger than their uncertainties. The horizontal dashed line is the quantum critical point ( $\lambda_c = 0.5$ ) in the thermodynamic limit. The other parameters are the same as in panel (a).

within a reasonably good accuracy, even for a large system size.

The slow growing rate of  $t_{\text{RSNN}}$  as a function of the system size  $N$  can be understood as follows: The preparation time  $t_S$  of each data sampling depends on the system size linearly for the calculation of matrix elements, while the training time  $t_{\text{train}}$  depends on the model parameters as well as training scheme only. These two timescales determine the time for data preparation but are not sensitive to the system size. That is why RSNN could be more efficient than other numerical methods, especially for a larger system size. In the inset of Fig. 2(b), we also show the finite-size scaling of the phase-transition

point, which is defined when the difference of the lowest two eigenenergies is larger than their uncertainty. We could find that the precise determination of phase-transition point is also possible due to the generation of a large amount of data within reasonable accuracy.

#### IV. ACCURACY AND COMPUTATION TIME FOR DIFFERENT HYPERPARAMETERS

From the basic calculation shown above for 1D IMTF, we have demonstrated the possibility to apply RSNN for the study of a quantum many-body system. However, we have to emphasize that the results obtained above are nontrivial, especially for a large system size, because the Hilbert space grows exponentially as  $N$  increases, and hence only an exponentially small fraction of matrix elements are included in the RSNN. Therefore, the success of RSNN here results from the fact that the obtained simulation function  $F_{\text{RSNN}}$  captures how the eigenvalues and/or other physical parameters change as a function of the system parameter  $\lambda$  through a small portion of matrix elements. To demonstrate this, below we systematically investigate how the accuracy and efficiency of RSNN can change by tuning the hyperparameters of RSNN during the training process.

In Fig. 3(a), we show how the average accuracy of eigenstate energies increases as the number  $N_S$  of sampling matrices increases. This reflects the fact that including more sampling matrices enhances the accuracy as the nature of neural networks. This also implies that such a high accuracy of prediction in the test regime could not be obtained faithfully if one just fits the eigenstate energy curves without knowing the matrix elements (i.e.,  $N_S \rightarrow 0$ ). This result, therefore, demonstrates the validity of our *random sampling hypothesis* as described in Eq. (2). Furthermore, we find that the average computation time  $t_{\text{RSNN}}$  does not increase but eventually becomes saturated for large  $N_S$  because we have applied the early stop method during the training process to avoid overfitting. In Fig. 3(b), we show the same calculation as a function of sampling matrix dimension  $M$  with  $N_S = 200$  being fixed. We also find that the computation time grows significantly in the small- $M$  regime but becomes saturated as  $M > 10$  due to the early stop mechanism. Note that, comparing with the calculation for larger values of  $N_S$  in panel (b), it requires much more computational memory in RSNN when the dimension  $M$  of sampling matrices increases, since the input features (matrix elements) scale as  $M^2$ . Therefore, here we just show the calculated results up to  $M = 20$  and expect the accuracy could grow further for a larger value of  $M$ .

Finally, in Fig. 3(c), we show how the accuracy and average computation time changes as a function of  $\lambda_0$ , which measures the relative size of training regimes by  $R_{\text{train}} \equiv (0, \lambda_0) \cup (1 - \lambda_0, 1)$ , see Fig. 2(a). Here we have fixed the total number  $N_{\text{train}}$  of the training data in the training regime with  $N_S = 200$  and  $M = 10$  for all different values of  $\lambda_0$ . As expected, the calculated results show that the overall accuracy of the RSNN prediction increases monotonically as a function of  $\lambda_0$  and reaches 100% when  $\lambda_0 \rightarrow 0.5$  because the test regime is so close to the training regime. On the other hand,  $t_{\text{RSNN}}$  still remains almost constant since the total number of training data is the same.

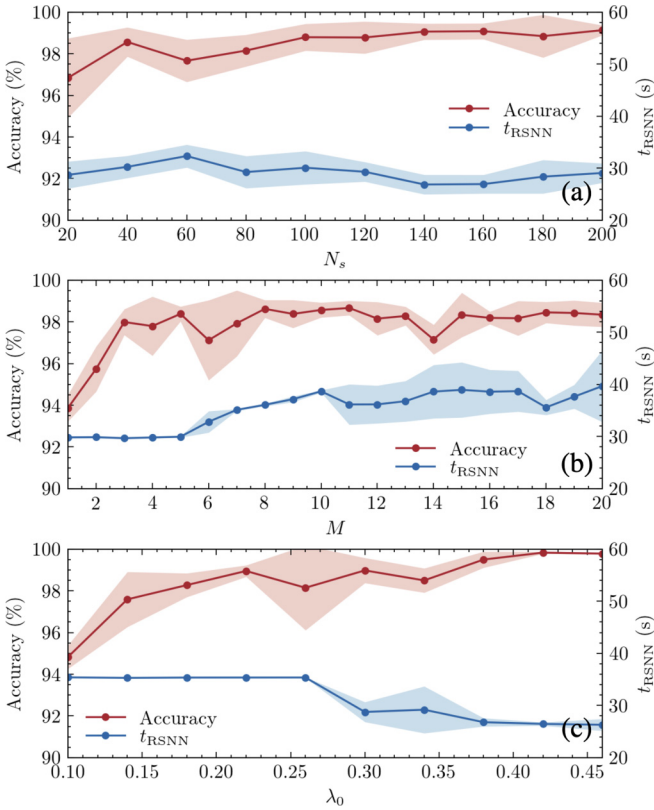


FIG. 3. (a) Accuracy and average computation time of RSNN ( $t_{\text{RSNN}}$ ) for the lowest three eigenenergies of 1D IMTF as a function of the number  $N_s$  of sampling at each  $\lambda$ . Here  $N = 28$ ,  $N_s = 200$ , and  $M = 10$ . Panel (b) shows the same calculation but for a different size of sampling matrix,  $M$ , for  $N_s = 50$  and  $N = 28$ . The average computation time stops increasing due to the early stop mechanism when the training loss is saturated. Panel (c) shows the same calculation for different value of training regime,  $\lambda_0$  (see the text). The total number of  $\lambda$ s for different training regimes are still the same ( $N_{\text{train}} = 10\,000$ ). Here  $N_s = 200$ ,  $M = 10$ , and  $N = 28$ .

## V. ENERGY SPECTRUM IN THE STRONGLY INTERACTING REGIME OF THE ONE-DIMENSIONAL FERMION-HUBBARD MODEL

The 1D IMTF discussed above is a good example for the application of RSNN, but it is not a strongly correlated system because its Hamiltonian can still be mapped to a free fermion model via the Jordan-Wigner transformation [30] and hence the eigenstates are still product states without correlation. To investigate the application of RSNN in the strongly correlated regime, here we consider the 1D Fermi-Hubbard model (1D FHM) with the following system Hamiltonian:

$$\hat{\mathcal{H}}_{\text{FH}} = -t \sum_{i,s} (\hat{c}_{i,s}^\dagger \hat{c}_{i+1,s} + \text{H.c.}) + U \sum_i \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow}, \quad (4)$$

where  $\hat{c}_{i,s}$  and  $\hat{n}_{i,s}$  are the fermion field operator and the number operator at site  $i$  and of spin  $s = \uparrow/\downarrow = \pm$ .  $t$  and  $U$  are the hopping energy and the on-site repulsion, respectively.

It is well known that, in the weakly interacting limit ( $U/t \rightarrow 0$ ), the 1D FHM can be well described by a Luttinger liquid [33], where all elementary excitations are bosonic and collective modes, and which can be separated into spin and

charge sectors (i.e., spin-charge separation). When considering the backward scattering for a large momentum transfer as well as the Umklapp scattering in the presence of a periodic lattice, the spin-charge excitations become gapless at the momentum  $p = 2k_F$  and  $4k_F$ , respectively, even for an infinitesimal  $U > 0$ . In the strongly interacting limit ( $U/t \rightarrow \infty$ ), on the other hand, states with different particle distributions are almost degenerate with either zero or one particle per site. The system then becomes equivalent to the  $t$ - $J$  model [34] with an antiferromagnetic spin-exchange coupling through the second-order perturbation of  $t$  (i.e.,  $J \propto t^2/U$ ).

Below we use 1D FHM as an example to test RSNN in the strongly correlated regime, after the model is trained in the weakly interacting regime. More precisely, we train a RSNN model by using the exact results of momentum-energy dispersion obtained by the Bethe-ansatz (BA) method [35,36]. The nonlinear coupled algebraic equations within the BA method is described by

$$e^{ik_j L} = \prod_{\alpha=1}^{N_\downarrow} \frac{\sin k_j - \lambda_\alpha + iU/4}{\sin k_j - \lambda_\alpha - iU/4}, \quad (5)$$

$$\prod_{j=1}^N \frac{\lambda_\alpha - \sin k_j + iU/4}{\lambda_\alpha - \sin k_j - iU/4} = -\prod_{\beta=1}^{N_\downarrow} \frac{\lambda_\alpha - \lambda_\beta + iU/2}{\lambda_\alpha - \lambda_\beta - iU/2}, \quad (6)$$

where  $L$  and  $N$  is the total number of sites or fermions and  $N_\uparrow$  ( $N_\downarrow$ ) is the number of spin-up (spin-down) fermions ( $N_\downarrow \leq N/2$ ). The pseudomomentum  $\{k_j\}$  and spin rapidities  $\{\lambda_\alpha\}$  are variables to be solved and are related to the total energy by  $E = -2t \sum_{j=1}^N \cos k_j$  and the total momentum by  $p = \sum_{j=1}^N k_j$ . For simplicity, here we just consider the energy spectrum of holon excitations in the charge sector [36]. Results of spinon excitations in the spin sector can also be obtained similarly.

In Figs. 4(a) and 4(b), we show the predicted energy spectrum ( $p_i, E_i$ ) obtained by the RSNN approach at the intermediate [ $\lambda \equiv U/(t+U) = 0.76$  or  $U/t = 3.3$ ] and strong interaction strength ( $\lambda = 0.93$  or  $U/t = 15.3$ ), respectively. The training regime is in the weakly interacting regime  $R_{\text{train}}(\lambda) = (0.23, 0.73)$  or  $R_{\text{train}}(U/t) = (0.3, 2.8)$ . The input features are  $16 \times 16$  random sampling matrices, obtained from the original system Hamiltonian (see Fig. 1 and Sec. II), and the output label is the whole energy spectrum, ( $p_i, E_i$ ). In the training regime, we take  $N_{\text{train}} = 2000$  values of  $\lambda$  and generate  $N_s = 100$  sampling matrices for each of them. Compared to the exact solution by BA, the predicted results of RSNN are pretty good even in the strongly interacting regime. In Fig. 4(c) we show the associated density of states (DOS), which reflects the interaction-broadened bandwidths.

In Fig. 4(d), we show the obtained accuracy and its uncertainty (obtained by averaging five independent calculations) for the whole test regime. We could find that the accuracy of the whole energy spectrum could be as high as 99% in the intermediate interaction regime ( $\lambda \sim 0.76$  or  $U/t \sim 3.3$ ), while it decreases gradually in the strongly interacting regime ( $\lambda > 0.8$  or  $U/t > 4$ ) with a larger uncertainty at the same time. However, we note that it is still impressive that the accuracy can be still larger than 95% even for  $\lambda = 0.97$  ( $U/t = 30.3$ ), showing that RSNN could be a very promising tool to estimate physical quantities in a strongly correlated system even it is trained in the weakly interacting regime.

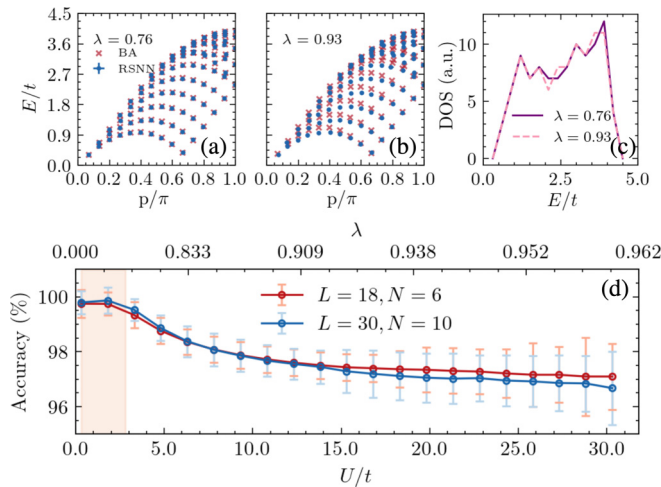


FIG. 4. Panels (a) and (b) show the predicted holon excitation spectrum (blue dots) of 1D FHM by RSNN for  $\lambda \equiv U/(t + U) = 0.76$  and  $0.93$ , respectively. The training regime is  $R_{\text{train}} = (0.23, 0.73)$  with  $L = 30$ ,  $N = 10$ , and  $N_{\downarrow} = 5$ . Exact results calculated by using the Bethe ansatz (BA, red crosses) are also shown together for comparison. Panel (c) shows the associate density of states (DOS) obtained by RSNN. Panel (d) shows the accuracy of the energy spectrum with an uncertainty by RSNN in the test regime,  $R_{\text{test}} = (0.73, 0.96)$  (white background). Results for two different system sizes are shown together for comparison.

## VI. QUANTUM CRITICAL EXPONENTS OF ONE-DIMENSIONAL XXZ MODEL

The quantum phase transition we discussed in the 1D IMTF is a first-order transition, where the magnetization changes discontinuously in the thermodynamic limit ( $N \rightarrow \infty$ ). It is therefore instructive to see if RSNN could be also applied to the study of the second-order phase transition, where the order parameter changes continuously in the thermodynamic limit and hence scaling exponents could be identified near the quantum critical point (QCP) [37].

One of the most important examples is the superfluid to Mott insulator transition for strongly interacting bosonic atoms loaded in an optical lattice [38]. Here we consider a simpler case of hard-core bosons with a finite intersite repulsion to compete with the kinetic energy, leading to the so-called Bose  $t$ - $V$  model:

$$\hat{\mathcal{H}}_{t-V} = \sum_{i=1}^N [-t(\hat{b}_i^\dagger \hat{b}_{i+1} + \text{H.c.}) + V \hat{n}_i \hat{n}_{i+1} - \mu \hat{n}_i], \quad (7)$$

where  $\hat{b}_i$  and  $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$  are the bosonic field operator and number operator, respectively,  $t$  and  $V$  are the tunneling and interaction between the nearest-neighbor sites, and  $\mu$  is the chemical potential. It is easy to see that the system prefers to be superfluid if  $V$  is small and can become a solid phase at half filling when  $V$  is repulsive and large. Quantum phase diagrams of such a superfluid-to-solid transition has been studied by quantum Monte Carlo methods in 1D and 2D systems [12].

Since the number of particles per site is either 0 or 1 in the hard-core limit, it is easy to connect such a  $t$ - $V$  model of hard-core bosons to a spin- $\frac{1}{2}$  system. More precisely, in the dilute limit, one could use Holstein and Primakoff transfor-

mation [39] to map the above  $t$ - $V$  model into the spin- $\frac{1}{2}$  XXZ model with a transverse field. To simplify the calculation in the rest of this paper, however, we concentrate on the quantum phase transition of the 1D XXZ model itself at zero field, which has the following system Hamiltonian:

$$\hat{\mathcal{H}}_{XXZ} = -\frac{J}{2} \sum_{j=1}^N (\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y + \lambda \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z), \quad (8)$$

where  $J$  is the in-plane spin coupling and  $\lambda$  is the  $z$ -direction spin coupling. It is well known that there are three different phases for 1D XXZ model in the thermodynamic limit: Antiferromagnetism (AFM, gapped) for  $\lambda < -1$ , paramagnetism (PM, gapless) for  $-1 < \lambda < 1$ , and ferromagnetism (FM, gapped) for  $\lambda > 1$ . The superfluid-to-solid transition of the  $t$ - $V$  model of hard-core bosons corresponds to the AFM-PM transition at  $\lambda = -1$ , which we study closely by RSNN here. We note that the spin  $\frac{1}{2}$  1D XXZ model in Eq. (8) could be also exactly solved by the Bethe-ansatz method [2,40,41].

To investigate the quantum phase-transition point near  $\lambda = -1$ , we use  $\lambda \in (-1.15, -1.125) \cup (-0.875, -0.85)$  as the training regime with  $N_{\text{train}} = 500$  for the training data inside. For each  $\lambda$ , we generate  $N_s = 100$  random sampling matrices (with the dimension  $M = 10$ ) as input features. In Fig. 5(a), we show the predicted spectrum for  $\lambda = -1.1$  (in the test regime) for  $N = 28$ . The obtained spectrum agrees very well with the BA results (not shown). The lowest-energy excitation occurs at  $p = \pi$  as expected. The average accuracy of the whole energy spectrum is  $98.01\% \pm 1.29\%$  in the whole test regime, showing a pretty good prediction even near the phase-transition point,  $\lambda = -1$ .

In Fig. 5(b) we show the calculated spinon excitation gap  $\Delta$  as a function of  $\lambda$  for various system sizes  $N$ . We find that the gap almost vanishes for  $\lambda > -1$  as  $N > 300$ . However, as the system size increases, the matrix elements in the sampling Hamiltonians cover an increasingly smaller fraction of the original Hilbert space and therefore the prediction accuracy also decreases down to  $90.49\% \pm 1.09\%$  for  $N = 300$ .

As opposed to the first-order phase transition of 1D IMTF, the order parameter (here the spinon excitation gap) should decrease to zero continuously at the QPT point  $\lambda_c$  in the thermodynamic limit, but it always has a finite value for finite  $N$ . To determine the QCP from finite-size scaling, here we use the phenomenological renormalization group (PRG) method [42,43]: Using the fact that the excitation gap must scale with  $N$  linearly at the QCP in a 1D system, i.e.,  $\Delta(N, \lambda_c) \propto N^{-1}$ , it is reasonable to expect that, for any two large system sizes,  $N \neq N' \gg 1$ ,  $\Phi(N, N', \lambda) \equiv |1 - N' \Delta(N', \lambda) / N \Delta(N, \lambda)|$  has a minimum at  $\lambda_{\min}(N, N')$ . This minimum value of  $\Phi$  should reach zero and  $\lambda_{\min}(N, N') \rightarrow \lambda_c$  when  $N, N' \rightarrow \infty$  at the same time. As a result, after considering the possible uncertainty of the finite-size calculation, we define  $\lambda_c^{\text{RBNN}}(N) \equiv \text{Ave}_{N'}[\lambda_{\min}(N, N')]$ , where  $\text{Ave}_{N'}[\dots]$  is the average of different system sizes  $N'$  by keeping another one ( $N$ ) fixed.

In Fig. 5(c) we show the gap as a function of  $1/N$  in a log-log plot with different values of  $\lambda$  near the quantum critical point (QCP),  $\lambda_c = -1$ . We see that the curves approach linear when  $\lambda$  is increased from below  $\lambda_c$  as expected. In the inset, we show the calculated function,  $\Phi(N, N', \lambda)$ , for  $N = 400$

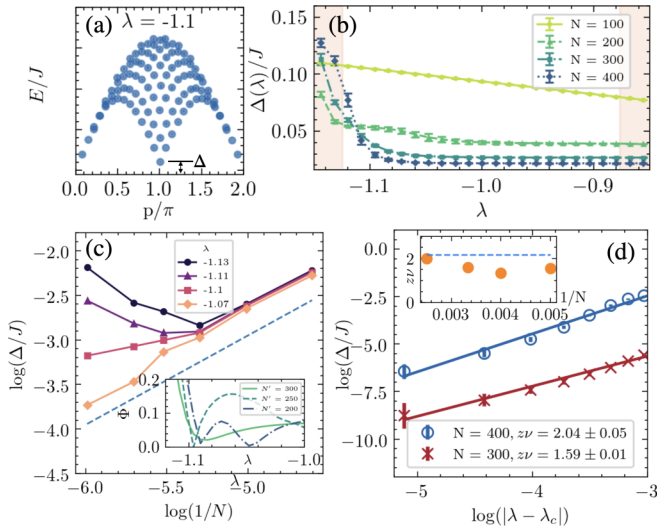


FIG. 5. (a) Spin excitation spectrum of 1D XXZ model for  $N = 28$  and  $\lambda = -1.1$ . The spin excitation gap  $\Delta$  is defined by the excitation energy at  $p = \pi$ . (b) The spin excitation gap  $\Delta$  predicted by RSNN for several system sizes up to  $N = 400$  with the corresponding uncertainties. The training and test regime is indicated by the colored and white background. (c) RSNN-predicted  $\Delta$  as a function of  $1/N$  (on a log-log plot) for different values of  $\lambda$ . The predicted phase-transition point,  $\lambda_c^{\text{RSNN}} = -1.07 \pm 0.02$ , is defined as the fixed point of phenomenological renormalization group (see text). The dashed line is the slope given by  $\lambda_c^{\text{RSNN}}$ . The inset shows the  $\Phi(\lambda)$  obtained (see text) for several system sizes. The position of its minimum value gives an estimate of critical quantum phase-transition point. (d) Predicted  $\Delta$  as a function of  $|\lambda - \lambda_c^{\text{RSNN}}|$  (on a log-log plot) for  $M = 300$  and  $400$ . The solid lines are the fitting results for different values of dynamical exponent. Inset shows how such critical exponent ( $z\nu$ ) changes as a function of the system size.  $z\nu = 2.16$  in the thermodynamic limit, as indicated by the horizontal dashed line.

and  $N' = 200, 250$ , and  $300$ , respectively. We could see that there are two local minimum values near  $-1.1$  and  $-1.05$ . After averaging the position of the minimum values, we obtain the estimated QCP at  $\lambda_c^{\text{RSNN}} = -1.07 \pm 0.02$ , pretty close to the value  $\lambda_c^{\text{BA}} = -1.06$  obtained by the finite-size Bethe-ansatz method. However, we have to emphasize that, for the results calculated by BA in the same regime (not shown here),  $\Phi(N, N', \lambda)$  is a pretty flat function with only one shallow minimum, different from the results predicted by RSNN. Therefore, what we could say is that the results predicted by RSNN here is just an estimate of the QCP, based on the data trained outside the critical regime.

Finally, using the obtained QCP value,  $\lambda_c^{\text{RSNN}} = -1.07$ , we could further calculate the critical exponent  $z\nu$ , which is defined by how the gap function vanishes [37] near the QCP:  $\Delta \sim |\lambda - \lambda_c|^{z\nu}$  for  $\lambda < \lambda_c$  in the thermodynamic limit (note that  $\Delta = 0$  for  $\lambda > \lambda_c$ ). In Fig. 5(d), we show that such a nontrivial scaling exponent could be obtained to be  $z\nu = 2.04 \pm 0.05$  by RSNN and is close (within 5% uncertainty) to the numerical value,  $z\nu = 2.16$ , obtained by the Bethe-ansatz in thermodynamics limit [3] (see Appendix B; note that the value of  $\nu$  is known to be one for the 1D XXZ model [44,45]). Again, we find RSNN could provide a reasonably good esti-

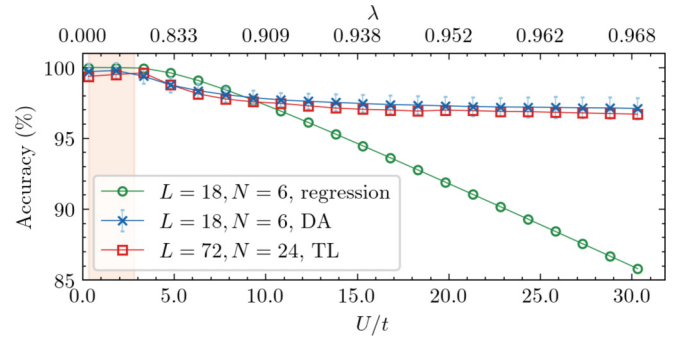


FIG. 6. Comparison of the accuracy for the energy spectrum of 1D FHM in different extended RSNN models. All the training data are taken from the weak-interaction regime (pink background,  $U/t < 3$ ) with  $L = 18$  and  $N = 6$ , while the test data are calculated in the medium- and strong-interaction regimes ( $3 < U/t < 30$ ). Green circles, blue crosses, and red boxes are respectively the results obtained by a naive linear regression method, by data augmentation (DA) during the training process, and by the transfer learning (TL) model for a larger system size with  $L = 72$  and  $N = 24$ . See the text for more details of these calculations.

mate of the critical exponent even if using the data outside the critical regime.

## VII. DISCUSSION

### A. Comparison between random sampling neural network and linear regression

Although we have demonstrated the effectiveness for the RSNN method to simulate the functional relationship between the sampled Hamiltonian matrix elements and the desired physical quantities, it is still instructive to compare the obtained results to those given by a naive linear regression method, which uses linear or other simpler functions with fewer parameters to fit the training data and then extrapolates the results in the test regime. In Fig. 6, we show the obtained accuracy of the regression result for 1D FHM with  $L = 18$  and  $N = 6$  (green circles). We find that, although the training data are well fit, the extrapolation accuracy falls off quickly as expected in the strong-interacting regime. On the other hand, as we have shown in Fig. 4(d), the accuracy of RSNN results could be kept larger than 97% even in the strongly interacting regime. As a result, the powerful simulation results of RSNN cannot simply be reproduced by a naive regression and extrapolation method.

### B. Data augmentation for fewer training data

In the applications of the RSNN method, we usually need many training data [ $N_{\text{train}} \sim O(10^3)$ ] to achieve satisfactory prediction results. It may be a serious drawback of RSNN because the computational time of these training data may be very long for a nonintegrable model without analytic solutions available. To overcome this problem, we could utilize the data augmentation (DA) method, which has been widely applied in computer vision or other machine-learning applications [46]. As an example to demonstrate DA, we again consider the energy spectrum prediction of 1D FHM by using five training

data only in the training regime. We then use linear interpolation to simulate other data points in the training model. As have been shown above, the interpolated results are very close to the exact solutions in the training regime and therefore could be used for the model training process.

However, different from the results of regression shown in the last section, we do NOT extrapolate these results in the test regime directly. On the other hand, the test data here are predicted by the RSNN model after using data augmentation in the training process. In Fig. 6, we find that the obtained accuracy (blue crosses) is almost the same ( $>97\%$ ) as the results obtained by using exact solutions as the training data [see Fig. 4(d)]. Therefore, by using such a simple data augmentation method, we could demonstrate that the time to generate training data could be strongly reduced (from 2000 to 5 in this case) almost without any loss of accuracy. Note that this result also applies to nonintegrable models because, from a machine-learning point of view, such a data-driven approach (such as RSNN) does not rely on any specific form of the Hamiltonian.

### C. Transfer learning for larger system sizes

In the theoretical studies of many-body physics, how to apply numerical methods to systems of large sizes in the thermodynamic limit is always a challenging problem because the required computational time and memory grows significantly. In the application of RSNN we proposed here, however, this situation could be handled differently by taking advantage of neural networks through the transfer learning (TL) method [47], which consists of two stages: We first use the abundant data of a smaller system to pretrain a RSNN model; second, we fix the parameters of the obtained RSNN model in the first few neural layers and then train the parameters in the last few layers by using fewer training data in a larger system. By taking this two-step training process, our RSNN model could then catch most of the important physics even when the available training data in a larger system size are much fewer.

As an example, in Fig. 6 we show TL results for the 1D FHM: The pretraining data are from a small system size with  $L = 18$ ,  $N = 6$ , and  $N_{\text{train}} = 1000$  in the weakly interacting regime, while we could obtain the results with a pretty high accuracy ( $\approx 97\%$ , see the red boxes) in the strongly interacting regime even for a larger system size ( $L = 72$  and  $N = 24$ ). Note that, in the second stage of training, we use only four training data to get the parameters in the last layer of RSNN to get this surprising result. Therefore, we demonstrate that our RSNN model could be easily extended to study a larger system size with a much fewer training data through the application of transfer learning method. This is a very unique advantage for the RSNN method in the application of many-body physics.

## VIII. SUMMARY

Motivated by the pattern-recognition method in computer vision, we propose an approach to predict the physical quantities of a general many-body system by randomly sampling the whole system Hamiltonian through a self-supervised learning

process. The training data could be obtained by perturbation theory or other existing numerical methods in the weakly interacting regime (or any certain parameter regimes). We have systematically investigated its applicability in several 1D exactly solvable models and demonstrate how it could provide pretty good prediction results of the ground-state energy, the momentum-energy spectrum, magnetization (or other order parameters), as well as the quantum phase-transition point and the associated critical exponents. One of the most important advantages of RSNN is that one just needs to train the model one time in the training regime and then gets an arbitrary amount of data immediately in the test regime, even in the strongly correlated regime or near the quantum phase-transition point. We further show that, by using data augmentation and transfer learning methods, one could apply RSNN easily to a larger system with much fewer training data, which could be provided more easily and quickly by existing numerical methods. Therefore, we believe that RSNN is a very effective and important approach to complement the existing numerical methods for exploring quantum many-body problems in a much wider parameter space.

We evaluate our models on Google Colaboratory cloud computing platform, which specifies two cores Intel(R) Xeon(R) CPU @ 2.20 GHz and NVIDIA Tesla P100 GPU. We provide Github code [49] for the application of RSNN in a 1D Ising model with a transverse field.

## ACKNOWLEDGMENTS

We thank Ming-Chiang Chung, Pochung Chen, Chung-Yu Mou, and Chung-Hou Chung for fruitful discussions. This work is supported by the Ministry of Science and Technology grant (MOST 107-2112-M-007-019-MY3) and by the Higher Education Sprout Project funded by the Ministry of Science and Technology and the Ministry of Education in Taiwan. We thank the National Center for Theoretical Sciences for providing full support.

## APPENDIX A: MODEL PARAMETERS OF RANDOM SAMPLING NEURAL NETWORK

In the construction of the RSNN, as introduced in Sec. II, we have two types of hyper-parameters: one is about the process for random sampling and the other is about the structure of the CNN structure, which accepts the input of random sampling matrices and outputs the expected physical quantities through neural networks (see also Fig. 1). Since we have mentioned the parameters used for random sampling in the text for each model (say, matrix size  $M$ , number of sampling matrices,  $N_S$ , number of training data,  $N_{\text{train}}$ , and training regime,  $R_{\text{train}}$ ), here we provide further information about the hyper-parameters used for the second part in the CNN structure.

As for the final output layer, we use a loss function to constrain the output to be our desired values for a self-supervised learning process. For example, if the output is to simulate the three lowest exactly known eigenenergies of the 1D IMTF system as shown in Sec. III, the loss function we used is



TABLE I. Hyper-parameters used for the CNN part of RSNN in the examples of this paper (see text).  $L_{\text{conv}}$  refers to the number of convolutional layers,  $L_{\text{fc}}$  refers to the number of fully connected layers, and  $N_{\text{neurons}}$  is the number of neurons for each layer, as shown in the list.

Case parameter	$L_{\text{conv}}$	$L_{\text{fc}}$	$N_{\text{neuron}}$
IMTF (eigenenergies)	2	2	[1250,10]
IMTF (magnetization)	2	2	[625,100]
FHM (holon spectrum)	2	6	[625,102,102,102,102,102]
XXZ (spinon spectrum)	2	6	[625,102,102,102,102,102]
XXZ (spinon gap)	2	3	[625,40,20]

designed as follows:

$$\text{Loss} = |\text{Pred} - \vec{E}^{\text{ED}}| + \beta \sum_i (W_i^2 + b_i^2), \quad (\text{A1})$$

where  $\text{Pred}$  is the predicted results of the neural network for each run and  $\vec{E}^{\text{ED}} \equiv [E_0^{\text{ED}}, E_1^{\text{ED}}, E_2^{\text{ED}}]$  are the lowest three eigenstate energies provided by exact diagonalization (or by results known before). The first term is taken as the batch average, and second term is to constrain the magnitudes of weighting ( $W_i$ ) and bias ( $b_i$ ) of all neurons (with index  $i$ ) from over-fitting.  $\beta > 0$  is an empirical parameter. The optimization process is done by Adam method [48] with the corresponding model parameters shown in Table I.

## APPENDIX B: EXPONENT $z\nu$ BY BETHE-ANSATZ

To extract the critical exponent of the 1D XXZ model, we calculate it from the Bethe-ansatz solution in the thermodynamic limit ( $L \rightarrow \infty$ ). For  $\lambda < -1$ , the dispersion relation in thermodynamics limit can be expressed in integral form as follows [3]:

$$E(P) = \frac{2K(m) \sinh \phi(\lambda)}{\pi} \sqrt{1 - m \sin^2 P}, \quad (\text{B1})$$

where  $E(P)$  is the energy dispersion of momentum  $P$ ,  $\lambda = -\cosh \phi$ ,  $\phi \equiv \pi K'(m)/K(m)$ , the parameter  $m = k^2$  with  $k$  being the elliptic modulus, and  $K(m)$  is the complete elliptic integral of first kind:

$$K(m) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - m \sin^2 \theta}}. \quad (\text{B2})$$

With a fixed  $\lambda$ , we can obtain the value of  $m$  by solving the differential equation  $\cosh^{-1}(-\lambda) = \pi K'(m)/K(m)$ , which evaluates the dispersion function (B1). Such a function exists at the lowest energy gap at  $P = \pi/2$ , thus the gap function  $\Delta(\lambda)$  is defined as

$$\Delta(\lambda) \equiv E(\pi/2) = \frac{2K(m) \sinh \phi(\lambda)}{\pi} \sqrt{1 - m}, \quad (\text{B3})$$

which has a fitting parameter (exponent)  $z\nu = 2.16$  for the form  $|\lambda - \lambda_c|^{z\nu}$  with  $\lambda_c = -1$ .

- 
- [1] P. Pfeuty, *Ann. Phys. (NY)* **57**, 79 (1970).  
[2] C. N. Yang and C. P. Yang, *Phys. Rev.* **150**, 321 (1966).  
[3] B. Sutherland, *Beautiful Models - 70 Years of Exactly Solved Quantum Many-Body Problems* (World Scientific Publishing Company, NJ, 2004).  
[4] N. Hugenholtz, *Physica* **23**, 481 (1957).  
[5] W. Tobocman, *Phys. Rev.* **107**, 203 (1957).  
[6] R. Shankar, *Rev. Mod. Phys.* **66**, 129 (1994).  
[7] K. G. Wilson, *Phys. Rev. B* **4**, 3174 (1971).  
[8] K. G. Wilson and J. Kogut, *Phys. Rep.* **12**, 75 (1974).  
[9] A. O. Gogolin, A. A. Nersisyan, and A. M. Tsvelik, *Bosonization and Strongly Correlated Systems* (Cambridge University Press, Cambridge, 2004).  
[10] J. von Delft and H. Schoeller, *Ann. Phys. (Berlin, Ger.)* **7**, 225 (1998).  
[11] R. R. dos Santos, *Braz. J. Phys.* **33**, 36 (2003).  
[12] M. Capello, F. Becca, M. Fabrizio, and S. Sorella, *Phys. Rev. Lett.* **99**, 056402 (2007).  
[13] W. M. C. Foulkes, L. Mitras, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).  
[14] U. Schollwöck, *Rev. Mod. Phys.* **77**, 259 (2005).  
[15] S. R. White and D. J. Scalapino, *Phys. Rev. Lett.* **80**, 1272 (1998).  
[16] R. Orus, *Nat. Rev. Phys.* **1**, 538 (2019).  
[17] A. W. Sandvik and G. Vidal, *Phys. Rev. Lett.* **99**, 220602 (2007).  
[18] S. Das Sarma, D.-L. Deng, and L.-M. Duan, *Phys. Today* **72**(3), 48 (2019).  
[19] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).  
[20] K. Choo, G. Carleo, N. Regnault, and T. Neupert, *Phys. Rev. Lett.* **121**, 167204 (2018).  
[21] B. Peter, C. Juan, R. G. Melko, and T. Simon, *Sci. Rep.* **7**, 8823 (2017).  
[22] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, *Phys. Rev. X* **7**, 031038 (2017).  
[23] Y. Ming, C.-T. Lin, S. Bartlett, and W.-W. Zhang, *npj Comput. Mater.* **5**, 88 (2019).  
[24] R. G. M. Juan Carrasquilla, *Nat. Phys.* **13**, 431 (2017).  
[25] X.-X. Niu and C. Y. Suen, *Pattern Recognit.* **45**, 1318 (2012).  
[26] H. Tang, M. Schrimpf, W. Lotter, C. Moerman, A. Paredes, J. Ortega Caro, W. Hardesty, D. Cox, and G. Kreiman, *Proc. Natl. Acad. Sci. USA* **115**, 8835 (2018).  
[27] C. Doersch, A. Gupta, and A. A. Efros, in *2015 IEEE International Conference on Computer Vision (ICCV)* (2015), pp. 1422–1430.  
[28] M. Baker and R. Patil, *Reliab. Comput.* **4**, 235 (1998).  
[29] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschnner, C. Becker, K. Sengstock, and C. Weitenberg, *Nat. Phys.* **15**, 917 (2019).  
[30] P. Jordan and E. P. Wigner, *Eur. Phys. J. A* **47**, 631 (1928).  
[31] M. L. Wall and L. D. Carr, *New J. Phys.* **14**, 125015 (2012).  
[32] D. Jaschke, M. L. Wall, and L. D. Carr, *Comput. Phys. Commun.* **225**, 59 (2018).  
[33] F. D. M. Haldane, *J. Phys. C: Solid State Phys.* **14**, 2585 (1981).

- [34] G. T. Zimanyi and E. Abrahams, *Phys. Rev. Lett.* **64**, 2719 (1990).
- [35] H. Bethe, *Eur. Phys. J. A* **71**, 205 (1931).
- [36] D. W. Wang and S. Das Sarma, *Phys. Rev. B* **65**, 035103 (2001).
- [37] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2011).
- [38] G. Markus, M. Olaf, E. Tilman, T. W. Hänsch, and B. Immanuel, *Nature (London)* **415**, 39 (2002).
- [39] T. Holstein and H. Primakoff, *Phys. Rev.* **58**, 1098 (1940).
- [40] C. N. Yang and C. P. Yang, *Phys. Rev.* **150**, 327 (1966).
- [41] C. N. Yang and C. P. Yang, *Phys. Rev.* **151**, 258 (1966).
- [42] T. Sakai and M. Takahashi, *J. Phys. Soc. Jpn.* **59**, 2688 (1990).
- [43] J. A. Plascak, W. Figueiredo, and B. C. S. Grandi, *Braz. J. Phys.* **29**, 579 (1999).
- [44] Klümper and Andreas, *Z. Phys. B: Condens. Matter* **91**, 507 (1993).
- [45] J. Suzuki, T. Nagao, and M. Wadati, *Int. J. Mod. Phys. B* **06**, 1119 (1992).
- [46] C. Shorten and T. Khoshgoftaar, *J. Big Data* **6**, 60 (2019).
- [47] K. Weiss, T. M. Khoshgoftaar, and D. Wang, *J. Big Data* **3**, 9 (2016).
- [48] D. P. Kingma and J. Ba, *3rd International Conference for Learning Representations*, San Diego (2015).
- [49] [https://github.com/CYLphysics/RSNN\\_TFIM1D](https://github.com/CYLphysics/RSNN_TFIM1D).