Multi-body wave function of ground and low-lying excited states using unornamented deep neural networks

Tomoya Naito (内藤智也) ⑩,[1,*] Hisashi Naito (内藤久資) ⑩,[2,†] and Koji Hashimoto (橋本幸士) ⑩[3,‡]

[1]*RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), Wako 351-0198, Japan*
and Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan
[2]*Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan*
[3]*Department of Physics, Kyoto University, Kyoto 606-8502, Japan*

We propose a method to calculate wave functions and energies not only of the ground state but also of low-lying excited states using a deep neural network and the unsupervised machine learning technique. For systems composed of identical particles, a simple method to perform symmetrization for bosonic systems and antisymmetrization for fermionic systems is also proposed.

I. INTRODUCTION

Atoms, molecules, and solids are composed of many electrons and ions, and atomic nuclei are composed of many nucleons. In principle, once the Schrödinger equation of these systems is solved, most properties can be described. However, in practice, they are quantum many-fermion systems, which are difficult to solve directly. Hence, it has been one of the important issues to solve the Schrödinger equation for the quantum many-fermion efficiently and accurately; in fact, many numerical methods, for instance, the Faddeev calculation [1], several methods for few-body systems [2–12], the quantum Monte Carlo method (QMC) including the variational Monte Carlo, diffusion Monte Carlo (DMC), and auxiliary-field quantum Monte Carlo methods [13–19], the configuration interaction method [20–22], the coupled cluster method [23–26], the density functional theory (DFT) [27–29], the density matrix renormalization group [30–33], the dynamical mean-field theory [34–36], and the lattice effective field theory [37–40] have been proposed in recent decades.

Among the above, DFT and QMC are classified into methods based on the variational principle. The variational principle [41] guarantees that the ground-state energy E_{gs} of a Hamiltonian H satisfies

$$E_{\mathrm{gs}} = \inf \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \tag{1}$$

where all the possible functions are considered in the infimum. The minimizer corresponds to the ground-state wave function. Indeed, it is scarcely possible to consider all the possible functions throughout minimizing the energy expectation value; hence, in practice, the calculation accuracy of a method based on the variational principle depends on the ansatz of a trial wave function. In other words, the size of the space of trial wave functions, in principle, determines the calculation accuracy.

For instance, a trial wave function of DFT is a Slater determinant, which is the simplest antisymmetric trial wave function. Owing to the simpleness of the ansatz, the numerical cost is drastically reduced, while it is known that interparticle correlation is partially missing [42].

In the QMC calculation, a Jastrow-type trial wave function [43] is often used. A Jastrow-type wave function $|\Psi\rangle$ consists of a single- (or sometimes multi-)Slater determinant $|\Phi_0\rangle$ and a correlation factor F, $|\Psi\rangle = F |\Phi_0\rangle$. With assuming F as a symmetric function, $|\Psi\rangle$ satisfies antisymmetry. Owing to the introduction of the factor F, interparticle correlations are described better than a single Slater determinant. Nevertheless, most QMC calculations optimize mainly F, while $|\Phi_0\rangle$ is optimized only around an initial ansatz [44]. In addition, an ansatz is introduced even for F; hence, accuracy also depends on the ansatz. Recently, based on the QMC calculation, a deep neural network (DNN) has been used for the ansatz of a trial wave function [45–47]. Since deep neural networks span much wider space, calculation accuracy is much improved, which is guaranteed by the universal approximation theorem [48,49]. Nevertheless, once the ansatz of a trial wave function is introduced, the systematic improvement of the calculation is difficult even with a DNN since the space trial wave functions span is already limited.

Another problem of variational-principle-based methods is calculation of excitation spectra. Excitation spectra are important quantities of molecules and atomic nuclei, while the variational principle [Eq. (1)] obtains the ground state only. Hence, another technique is needed to calculate excited states based on the variational principle. Indeed, a method to calculate low-lying excited states using the DMC calculation was proposed [50] by considering the orthogonality of wave

functions, while its application has been still limited [51], whereas calculation of excited states on top of the DFT ground state has been widely performed by using the random phase approximation or some other techniques [52–56], while there is still room to be improved. Recently, low-lying excited states were also obtained in Ref. [57] using the combination of the DNN, QMC, and the orthogonal condition.

In this paper, we propose a new method to calculate energies and wave functions of the ground state and low-lying excited states based on the variational principle. The ground-state wave function is assumed to be a DNN, which, in principle, is able to represent any function [48,49]. Using an essence of the machine learning technique—the minimization of the loss function—they are directly optimized by using the machine learning technique. In the machine learning technique, a network structure is introduced and the parameters of the network are optimized with minimizing the loss function. This process is often called training. In many works, these parameters are trained by using the given data called learning data, whose size is often large. On the contrary, the machine learning trained without learning data, called unsupervised machine learning, has also been used in many different fields. References [58,59] also assumed the ground-state wave function as a deep neural network that were optimized by using the machine learning technique: Ref. [58] performed calculation of few-body bosonic systems and Ref. [59] performed calculation of the simplest realistic system—a deuteron. Although these papers are pioneering works of unsupervised machine learnings for quantum many-body problems, the fermion antisymmetrization was not considered and excited states were not studied, while both ground and excited states of many-fermionic systems are interesting in general. Recently, Ref. [60] proposed a method to obtain the ground state of the many-body Schrödinger equation for fermionic systems by using a tensor neural network, while its implementation is involved and the antisymmetrization is, indeed, not perfectly guaranteed.

In this paper, on top of the method in Refs. [58,59], a simple method of the antisymmetrization for many-fermion systems or the symmetrization for many-boson systems is introduced. Then, low-lying excited states are sequentially calculated by using the orthogonality conditions and the variational principles. In this method, there is no need to discover a DNN architecture to generate (anti)symmetric wave functions. The (anti)symmetrization is put at the level of loss functions. Furthermore, the symmetrization and the antisymmetrization are implemented in almost the same way and the wave function perfectly satisfies (anti)symmetry. Thanks to the simplicity of the implementation, the numerical cost is quite small. We show that our method works successfully for popular examples in bosonic and fermionic quantum mechanical systems, providing a fundamental basis of the DNN method for quantum mechanics.

This paper is organized as follows: Section II is devoted to calculation of the ground state. The novel (anti)symmetrization is introduced. Section III is devoted to the calculation of low-lying excited states. All the calculations are performed in a MacBook Pro with the Apple M1 chip [MacBook Pro (13-inch, M1, 2020): MacBookPro17,1] and 16 GB memory. Section IV gives a summary of this paper.
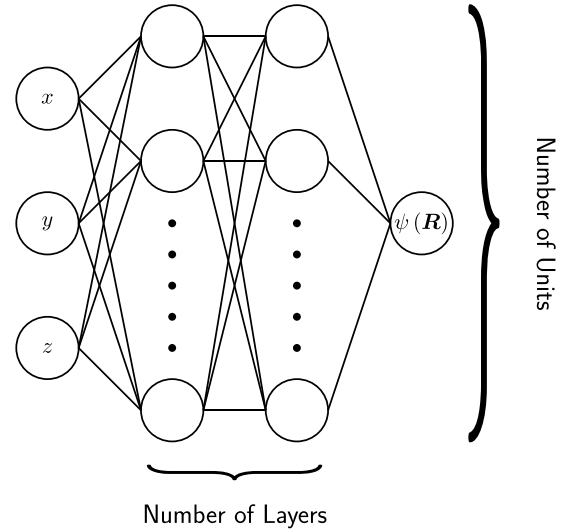

FIG. 1. Schematic figure of the deep neural network representing a one-dimensional three-body system.

II. GROUND-STATE CALCULATION

In this section, the ground-state wave function and energy are calculated by using a DNN and the machine learning technique. Throughout the paper, a machine learning software named TENSORFLOW [61] is used.

A. Network structure and machine learning technique

In general, a wave function of a d-dimensional N-particle system is a function of the spatial coordinates of all the particles $\boldsymbol{r}_j = (r_{j1}, r_{j2}, \ldots, r_{jd})$ ($j = 1, 2, \ldots, N$). Here, for the sake of simplicity, we neglect the spin and isospin dependence of wave functions and $\boldsymbol{R}$ denotes $\boldsymbol{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = (r_{11}, r_{12}, \ldots, r_{1d}, r_{21}, r_{22}, \ldots, r_{2d}, \ldots, r_{N1}, r_{N2}, \ldots, r_{Nd})$.

In this work, the wave function is represented by a deep neural network with Nd-input units that corresponds to the spatial coordinate $\boldsymbol{R}$ and one-output unit that corresponds to the value of the wave function $\psi(\boldsymbol{R})$. Between the input and output layers, there are hidden layers. Each unit is connected to all the units just one before or after layers. The schematic figure of the deep neural network is shown in Fig. 1. In this paper, the "softplus" function

$$\mathrm{softplus}(x) = \log(1 + e^x) \qquad (2)$$

is used for an activation function and the Adam optimizer [62] is used for the optimization process.

As the normal procedure of the numerical calculation of the wave function, the spatial coordinate is discretized. Each point is treated as a batch of the machine learning. In other words, if the spatial coordinate of each direction is discretized with M meshes, the batch size is the same as the number of meshes, M^{dN}. The mini-batch technique is not used.

Once the spatial coordinates are discretized, the Hamiltonian

$$H = -\frac{\hbar^2}{2m} \sum_j \Delta_j + \sum_j V^{\mathrm{ext}}(\boldsymbol{r}_j) + \frac{1}{2} \sum_{j \neq k} V^{\mathrm{int}}(\boldsymbol{r}_j, \boldsymbol{r}_k) \quad (3)$$

can be written as a matrix, where m is the mass of the particles, V^{ext} is the external potential, and V^{int} is the interparticle interaction. The matrices of the external potential and the interaction are diagonal and that of the kinetic energy is sparse. Hence, the expectation value of the Hamiltonian $\langle H \rangle$ can be calculated by using the sparse-matrix technique. The ground-state wave function minimizes $\langle H \rangle$; therefore, $\langle H \rangle$ is regarded as a loss function. Note that all the calculations are performed with double precision floating point numbers (`float64`). For simplicity, $m = \hbar = 1$ is assumed.

The procedure in the TENSORFLOW code is as follows:

(1) Construct a model of the deep neural network;

(2) Note that although a TENSORFLOW subroutine specification for the loss function technically requires two inputs—the training data (`true_value`) and the network output (`predicts`), the former is not referred in our training;

(3) Fit the model (`model.fit`) where the initial value of `predicts` consists of positive random numbers;

(4) The final wave function `output_wf` is given by using `model.predict`;

(5) The ground-state energy is calculated using the wave function obtained by the last step.

The third step (`model.fit`) corresponds to determining the parameters inside the DNN; the fourth step (`model.predict`) corresponds to storing the wave function obtained in the previous step; the fifth step corresponds to calculating the ground-state energy using the wave function obtained in the fourth step. Note that `predicts` and the final wave function should be normalized whenever generated.

B. One-dimensional one-particle systems

In this section, benchmark calculations of one-dimensional systems are shown. The dependence of the numbers of units and layers on calculation accuracy is also discussed. Since there exists only one particle, there is no interaction, $V^{\text{int}} \equiv 0$; thus, the Hamiltonian reads

$$H = -\frac{1}{2}\frac{d^2}{dx^2} + V^{\text{ext}}(x). \tag{4}$$

Since we focus only on bound states in this paper, it is enough to deal with the limited spatial region. In the calculation, $|x| \leqslant x_{\max}$ is considered and the box is discretized within 1024 meshes, i.e., $M = 1024$. The Dirichlet boundary condition $[\psi(\pm x_{\max}) = 0]$ is used. For the second derivative, the three-point derivative is used for simplicity, while it can be straightforwardly improved for the accuracy [63]. Then, H is discretized as

$$H \simeq \tilde{H} = -\frac{1}{2h^2}\tilde{T} + \tilde{V}^{\text{ext}}, \tag{5a}$$

$$\tilde{T} = \begin{pmatrix} -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -2 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -2 \end{pmatrix}, \tag{5b}$$

$$\tilde{V}^{\text{ext}} = \begin{pmatrix} V_1^{\text{ext}} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & V_2^{\text{ext}} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & V_3^{\text{ext}} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & V_{M-3}^{\text{ext}} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & V_{M-2}^{\text{ext}} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & V_{M-1}^{\text{ext}} \end{pmatrix}, \tag{5c}$$

and the wave function is also discretized as a $(M-1)$-dimensional vector

$$\psi \simeq \tilde{\psi} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_{M-3} \\ \psi_{M-2} \\ \psi_{M-1} \end{pmatrix}, \tag{6}$$

where $\tilde{\psi}$ is assumed to be normalized, i.e., $h\sqrt{\sum_j \tilde{\psi}_j^2} = 1$, $V_j^{\text{ext}} = V^{\text{ext}}(x_j)$, $\psi_j = \psi(x_j)$, $x_j = -x_{\max} + hj$, and h denotes the mesh size $h = 2x_{\max}/M$ [64]. This $\tilde{\psi}$ is used for `predicts` and `output_wf`. Here, a tilde denotes discretized form. Then, $\langle H \rangle$ can be calculated as

$$\langle H \rangle \simeq \langle \tilde{H} \rangle = \tilde{\psi}^{\mathsf{T}} \tilde{H} \tilde{\psi}. \tag{7}$$

1. Harmonic oscillator

First of all, the harmonic oscillator potential

$$V^{\text{ext}}(x) = \tfrac{1}{2}\omega^2 x^2 \tag{8}$$

is tested. The ground-state wave function ψ_{gs} and energy E_{gs} are, respectively, known exactly as [41]

$$\psi_{\text{gs}}(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \exp\left(-\frac{\omega x^2}{2}\right), \tag{9a}$$

$$E_{\text{gs}} = \frac{\omega}{2}. \tag{9b}$$

In this calculation, $x_{\max} = 5$ is used.

Table I shows the summary of calculations. In general, all the calculations give almost the correct energy [Eq. (9b)]. On the one hand, total optimization costs similar amount of time in all the calculation. On the other hand, different setup requires different number of epochs and time per epoch for optimizing the DNN. Small DNN tends to take shorter time for each epoch, while it requires longer epochs. It seems that 32 units per layer is too large, so it requires longer epoch and longer estimation time per epoch. It should be noted that the number of epochs differs in each run since the initial condition of the fitting procedure is generated by the random numbers. In addition, if one uses a different value of the learning rate, the number of epochs can be different.

TABLE I. Calculation summary of a one-body problem under the harmonic oscillator potential. Row with "—" in the column "# of unit for 2nd layer" corresponds to calculation performed only with one layer.

	# of Unit		Energy				
ω	1st Layer	2nd Layer	Kinetic	Potential	Total	# of Epochs	Time per Epoch (μs)
1.0	4	—	+0.250043	+0.250032	+0.500075	22 448	474.909
1.0	4	4	+0.250006	+0.249996	+0.500002	23 242	542.892
1.0	4	8	+0.250001	+0.249997	+0.499998	23 408	588.101
1.0	8	—	+0.250002	+0.250002	+0.500004	35 973	511.120
1.0	8	4	+0.250000	+0.249998	+0.499998	19 072	592.471
1.0	8	8	+0.250004	+0.249996	+0.500001	20 531	657.621
1.0	8	16	+0.249999	+0.249999	+0.499997	16 540	725.032
1.0	16	—	+0.250000	+0.249999	+0.499999	19 239	566.930
1.0	16	8	+0.250000	+0.249998	+0.499998	17 425	724.952
1.0	16	16	+0.249999	+0.249998	+0.499997	15 275	869.706
1.0	32	—	+0.250000	+0.249999	+0.499998	17 924	687.738
1.0	32	16	+0.249999	+0.249999	+0.499998	13 761	999.868
5.0	4	—	+1.250192	+1.250116	+2.500308	24 517	458.700
5.0	4	4	+1.250025	+1.249921	+2.499946	19 772	551.123
5.0	4	8	+1.250027	+1.249923	+2.499951	19 855	605.380
5.0	8	—	+1.249994	+1.249946	+2.499939	27 232	514.425
5.0	8	4	+1.250453	+1.249510	+2.499963	26 324	630.199
5.0	8	8	+1.249994	+1.249942	+2.499936	19 440	661.411
5.0	8	16	+1.250002	+1.249937	+2.499939	14 364	731.111
5.0	16	—	+1.249967	+1.249975	+2.499942	17 423	568.478
5.0	16	8	+1.250041	+1.249938	+2.499979	28 345	731.765
5.0	16	16	+1.249977	+1.249957	+2.499934	13 851	855.756
5.0	32	—	+1.249919	+1.250026	+2.499945	13 461	671.493
5.0	32	16	+1.249972	+1.249966	+2.499938	11 745	1005
10.0	4	—	+2.499927	+2.499866	+4.999793	21 134	482.504
10.0	4	4	+2.500071	+2.499694	+4.999765	23 623	554.103
10.0	4	8	+2.500000	+2.499720	+4.999719	24 949	605.551
10.0	8	—	+2.499877	+2.499945	+4.999822	16 636	513.759
10.0	8	4	+2.500038	+2.499718	+4.999756	20 628	598.401
10.0	8	8	+2.500003	+2.499730	+4.999733	19 261	664.986
10.0	8	16	+2.499947	+2.499770	+4.999717	15 053	731.275
10.0	16	—	+2.499873	+2.499896	+4.999769	17 351	561.501
10.0	16	8	+2.499996	+2.499759	+4.999756	14 979	730.518
10.0	16	16	+2.499917	+2.499822	+4.999739	15 437	859.019
10.0	32	—	+2.499834	+2.499885	+4.999719	16 520	671.640
10.0	32	16	+2.499976	+2.499745	+4.999721	19 833	996.732

Figure 2 shows relative errors of the loss function, $\langle H \rangle$, to the exact ground-state energy E_{gs} as functions of the number of epochs. It can be seen that, although the loss function achieved the relative error of 1.0×10^{-8}, the final accuracy becomes about 1.0×10^{-4}. This may be due to the precision of the TENSORFLOW code.
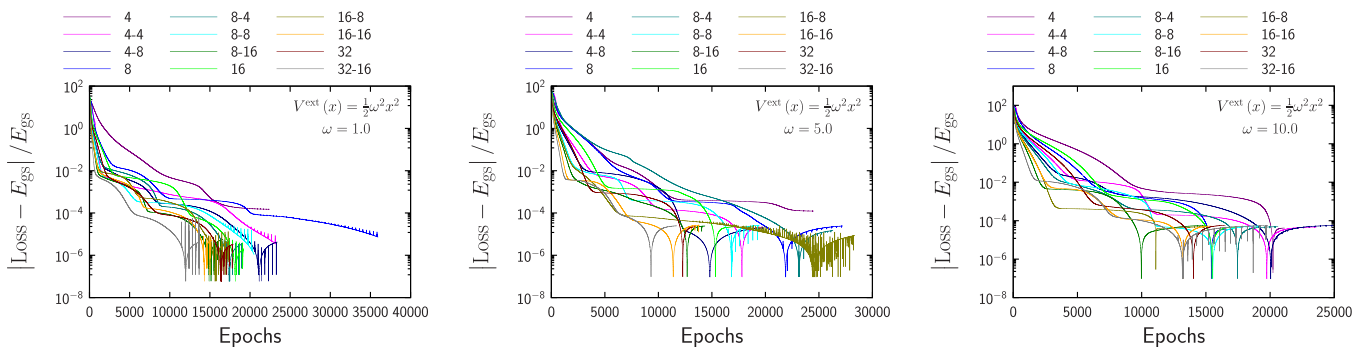

FIG. 2. Relative error of $\langle H \rangle$ to the exact ground-state energy E_{gs} for the harmonic oscillator potential as functions of the number of epochs.
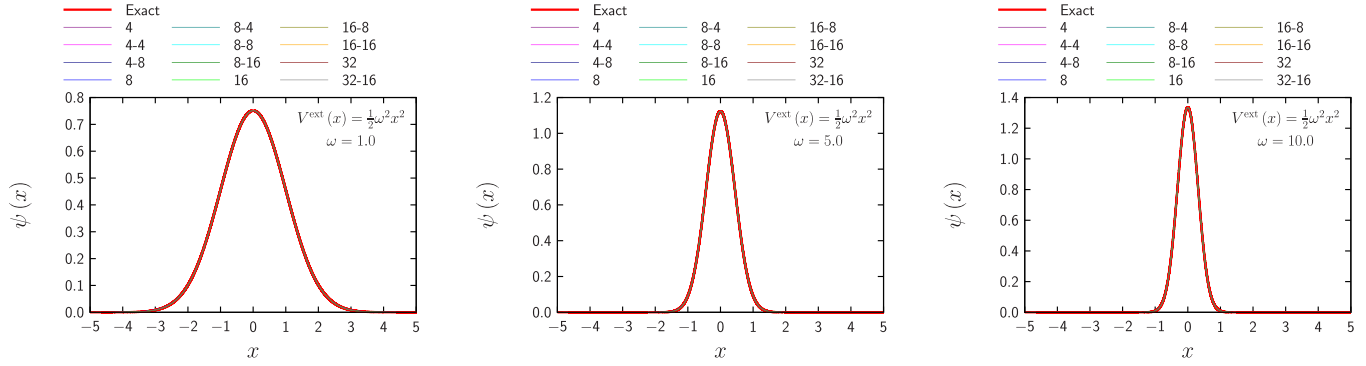
FIG. 3. Wave function under the harmonic oscillator potential. The red thick line corresponds to the exact solution [Eq. (9a)], while thin lines correspond to results of DNN calculation. Different thin line corresponds to different number of units. We observe that all the simulated results overlap with the exact solution.

Figure 3 shows calculated wave functions. The red thick lines correspond to the exact solution given in Eq. (9a), while thin lines correspond to the results given in this work, where different colors correspond to different numbers of units and layers. The relative errors of the DNN wave function, ψ^{DNN}, to the exact one, ψ^{exact},

$$\delta\psi(x) = \frac{\left|\psi^{\text{DNN}}(x) - \psi^{\text{exact}}(x)\right|}{\psi^{\text{exact}}(x)}, \tag{10}$$

are shown in Fig. 4. It can be seen that the DNN calculation, basically, reproduces the exact solution in our interest within the accuracy of 10^{-4} or more. This deviation can be reduced if we use more tight convergence criterion [65]. In the tail region, the deviation $\delta\psi(x)$ diverges, while this is because the denominator of Eq. (10), $\psi^{\text{exact}}(x)$, reaches to zero. The deviation looks larger if ω is smaller, which is related to the cutoff parameter for the spatial mesh $x_{\min}$. The exact value of $\psi_{\text{gs}}(x_{\min})$ is 2.8×10^{-6} for $\omega = 1.0$, while it is 8.1×10^{-28} for $\omega = 5.0$ and it is much smaller for $\omega = 10.0$, while in the numerical calculation they are approximated to zero. The value 2.8×10^{-6} may be too large to assume to be zero.

It should be noted that rather small DNN is enough to reproduce the solution of the wave function. Owing to the simplicity, it is easy to analyze the weights and biases of the DNN. For instance, the DNN wave function for the single layer with four units includes only 13 parameters; the ground-state DNN

wave function for $\omega = 1.0$ can be written as

$$\psi_{\text{gs}}(x) = \frac{1}{3.7451}\text{softplus}(a_{\text{gs}}(x)), \tag{11a}$$

$$a_{\text{gs}}(x) = 2.4069a_1(x) - 1.8344a_2(x) - 1.9778a_3(x)$$
$$+ 2.3484a_4(x) - 4.8998, \tag{11b}$$

$$a_1(x) = \text{softplus}(0.35953x + 3.9226), \tag{11c}$$

$$a_2(x) = \text{softplus}(2.5821x + 0.033213), \tag{11d}$$

$$a_3(x) = \text{softplus}(-0.65170x + 2.9574), \tag{11e}$$

$$a_4(x) = \text{softplus}(0.15421x + 2.2016), \tag{11f}$$

where the first coefficient of Eq. (11a) $(1/3.7451)$ is not obtained by the DNN but instead by the normalization. Hence, smaller DNN is better not only due to the calculation cost but also for analysis of the structure of DNN.

Let us provide our interpretation of the obtained wave function [Eq. (11a)]. The rectified linear function (ReLU)

$$\text{ReLU}(x) = \begin{cases} 0 & (x < 0), \\ x & (x \geqslant 0) \end{cases} \tag{12}$$

is a widely used activation function, and the softplus function can be regarded as a smoothed version of the ReLU. Here, for interpreting Eq. (11a), we shall just replace the softplus function with the ReLU. The wave function obtained by the DNN [Eq. (11a)] and the obtained function before the output layer [Eq. (11b)] are shown in Fig. 5, where the normalization
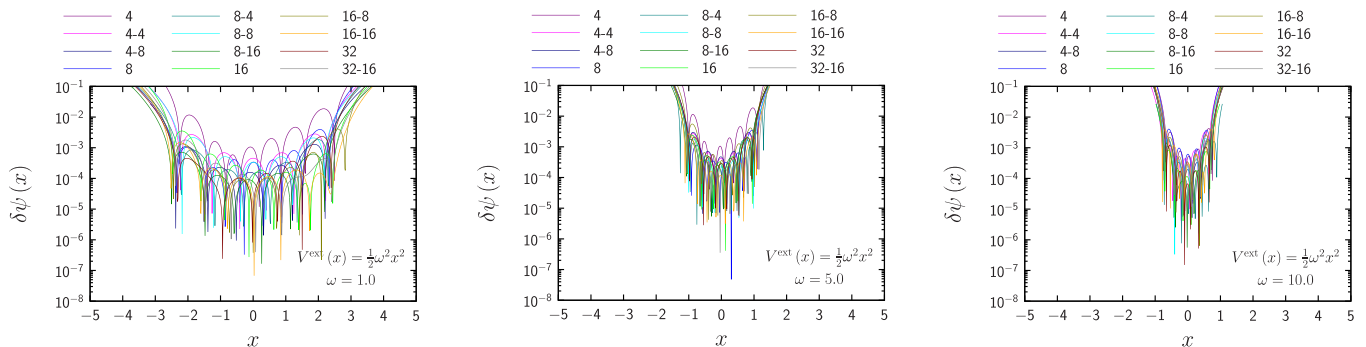

FIG. 4. Relative error of DNN wave function to exact one for the harmonic oscillator potential. Different thin line corresponds to different number of units. In the region of large $|x|$, the deviation $\delta\psi(x)$ diverges, because the denominator of Eq. (10), $\psi^{\text{exact}}(x)$, reaches to zero. Hence, the figures plot only $|\delta\varphi(x)| < 10^{-1}$.
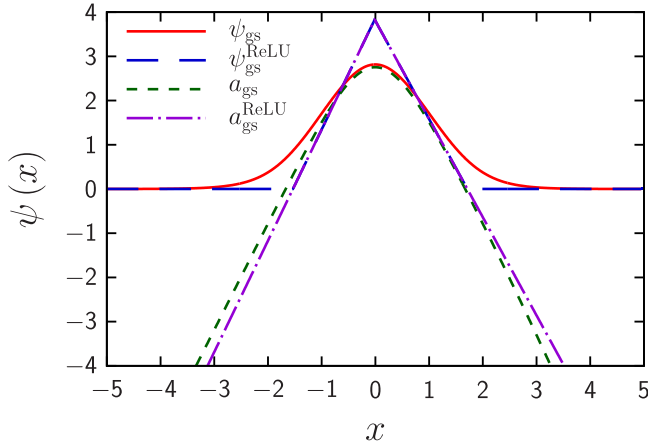
FIG. 5. Wave function obtained by the DNN [Eq. (11a)] and the obtained function before the output layer [Eq. (11b)], where the normalization is ignored. Equations (11a) and (11b) where the softplus function is replaced with the ReLU function are also plotted as $\psi_{\mathrm{gs}}^{\mathrm{ReLU}}$ and $a_{\mathrm{gs}}^{\mathrm{ReLU}}$, respectively.

factor $(1/3.7451)$ of ψ_{gs} is ignored. Equations (11a) and (11b) where the softplus function is replaced to the ReLU function are also plotted as $\psi_{\mathrm{gs}}^{\mathrm{ReLU}}$ and $a_{\mathrm{gs}}^{\mathrm{ReLU}}$, respectively. The DNN with the ReLU function can be understood as an approximation with a piecewise linear function. As shown in Fig. 5, the ground-state wave function in DNN is approximated by the following function:

$$\psi_{\mathrm{gs}} \approx \begin{cases} -ax + b & (0 \leqslant x \leqslant b/a), \\ ax + b & (-b/a \leqslant x \leqslant 0), \\ 0 & (\text{otherwise}), \end{cases} \qquad (13)$$

where a and b are positive numbers. The ReLU function at the output layer guarantees to make the wave function vanish for $x < -b/a$ and $x > b/a$, and thus, a_{gs} should be $\pm ax + b$. This function can be represented by just two ReLU functions. Hence, even two units in hidden layer are enough to describe the brief structure of ground-state wave function, and increasing the number of units, the ground-state wave function is reproduced easily. Since the ReLU function is not differentiable at $x = 0$, the ReLU wave function is not differentiable. Hence, the softplus is better to describe a differentiable function, while the ReLU function can also describe a differentiable function approximately if the number of units is large enough. In case of N-bodies systems, the similar function to Eq. (11a) can be represented by just 2^N ReLU functions.

2. Square-well potential

Next, the square-well potential

$$V^{\mathrm{ext}}(x) = \begin{cases} -V_0 & (|x| < x_0), \\ 0 & (\text{otherwise}) \end{cases} \qquad (14)$$

is tested ($V_0 > 0$). The analytical forms of the ground-state wave function ψ_{gs} and energy E_{gs} are unknown; thus, our values of the energy will be compared with the numerical calculation obtained by the orthodox method of Hamiltonian diagonalization. In this calculation, $x_{\max} = 20$ and $x_0 = 1$ is used.

Table II shows the summary of calculations. In general, all the calculations give almost the correct energy. The calculation time per epoch and the number of epochs with respect to the number of layers and units is slightly longer than the case of the harmonic oscillator.

Figure 6 shows relative errors of the loss function, $\langle H \rangle$, to the exact ground-state energy E_{gs} as functions of the number of epochs. It can be seen that, although the loss function achieved the relative error of 1.0×10^{-7}, the final accuracy becomes about 1.0×10^{-2}. Note that in this calculation, the convergence criteria is needed to be set looser than the other case; otherwise, it could not reach convergence. This may be related to the shape of the potential: Asymptotic region of the square-well potential is zero, while the harmonic oscillator potential increases rapidly. It will be shown later that the double-well potential, which is close to the latter situation, reaches convergence with the tight criterion.

Figure 7 shows calculated wave functions. The red thick lines correspond to the exact solution given by the exact diagonalization, where the same mesh size matrix form are used, for comparison; thin lines correspond to the results given in this work. It can be seen that the DNN calculation, basically, reproduces the solutions given by the exact diagonalization.

C. One-dimensional many-particle systems

When one considers systems composed of many identical particles, the symmetrization for bosonic systems or the antisymmetrization for fermionic systems of the wave function must be considered. The ground state of the bosonic system is identical to that of the different particles; hence it has no extra difficulty as was done in Ref. [58], while the antisymmetrization is rather difficult. In this section, a simple method of (anti)symmetrization in the DNN wave function is provided, in which the symmetrization and the antisymmetrization can be performed with equal footing.

1. Hamiltonian matrix

As was done in the last section, the discretized Hamiltonian $\tilde{H}$ should be represented in a matrix form and the discretized wave function $\tilde{\psi}$ should be represented in a vector form. Here, one-dimensional two-body systems are considered as an example, and their coordinates are denoted by x and y. Each direction is discretized with M meshes, i.e., in total $M \times M$ meshes. Then, the discretized wave function $\tilde{\psi}$ is

$$\tilde{\psi} = \begin{pmatrix} \psi_{11} \\ \psi_{12} \\ \vdots \\ \psi_{1(M-1)} \\ \psi_{21} \\ \psi_{22} \\ \vdots \\ \psi_{2(M-1)} \\ \vdots \\ \psi_{(M-1)1} \\ \psi_{(M-1)2} \\ \vdots \\ \psi_{(M-1)(M-1)} \end{pmatrix}, \qquad (15)$$

TABLE II. Calculation summary of an one-body problem under the square-well potential. Row with "—" in the column "# of unit for 2nd layer" corresponds to calculation performed only with one layer.

V_0	# of Unit		Energy			# of Epochs	Time per Epoch (µs)
	1st Layer	2nd Layer	Kinetic	Potential	Total		
0.5	4	—	+0.109852	−0.337864	−0.228012	26 600	579.787
0.5	4	4	+0.109935	−0.338093	−0.228159	44 065	717.730
0.5	4	8	+0.109944	−0.338110	−0.228166	27 771	778.485
0.5	8	—	+0.109852	−0.337876	−0.228024	21 802	619.806
0.5	8	4	+0.109962	−0.338088	−0.228126	37 964	776.147
0.5	8	8	+0.110009	−0.338173	−0.228164	36 410	856.319
0.5	8	16	+0.109968	−0.338119	−0.228151	69 731	968.769
0.5	16	—	+0.109923	−0.338026	−0.228104	17 015	734.647
0.5	16	8	+0.109967	−0.338116	−0.228148	22 157	1005
0.5	16	16	+0.109983	−0.338146	−0.228163	32 462	1155
0.5	32	—	+0.109847	−0.337872	−0.228025	24 879	921.208
0.5	32	16	+0.109976	−0.338136	−0.228160	34 549	1400
1.0	4	—	+0.206715	−0.812787	−0.606072	35 062	572.641
1.0	4	4	+0.206882	−0.813194	−0.606312	24 588	720.511
1.0	4	8	+0.206896	−0.813222	−0.606326	21 877	787.156
1.0	8	—	+0.206844	−0.813106	−0.606261	45 145	609.218
1.0	8	4	+0.206877	−0.813182	−0.606305	42 888	769.973
1.0	8	8	+0.207083	−0.813407	−0.606325	26 445	834.225
1.0	8	16	+0.206900	−0.813220	−0.606320	27 073	976.154
1.0	16	—	+0.206848	−0.813109	−0.606261	41 642	735.100
1.0	16	8	+0.206907	−0.813235	−0.606328	19 472	1002
1.0	16	16	+0.206888	−0.813202	−0.606314	25 746	1155
1.0	32	—	+0.206877	−0.813112	−0.606235	34 259	916.159
1.0	32	16	+0.206872	−0.813146	−0.606273	11 987	1410
5.0	4	—	+0.521370	−4.821182	−4.299812	45 997	565.210
5.0	4	4	+0.521033	−4.823422	−4.302390	34 568	722.257
5.0	4	8	+0.520966	−4.823344	−4.302378	38 875	774.016
5.0	8	—	+0.520887	−4.822751	−4.301865	32 216	610.063
5.0	8	4	+0.520913	−4.823243	−4.302329	17 236	766.967
5.0	8	8	+0.520972	−4.823484	−4.302512	23 526	829.876
5.0	8	16	+0.520893	−4.823451	−4.302558	22 945	971.763
5.0	16	—	+0.520781	−4.822461	−4.301680	27 332	740.434
5.0	16	8	+0.521167	−4.823316	−4.302150	15 343	995.692
5.0	16	16	+0.520927	−4.823470	−4.302542	18 961	1151
5.0	32	—	+0.520812	−4.822494	−4.301681	30 846	917.315
5.0	32	16	+0.520893	−4.823362	−4.302470	15 731	1413
10.0	4	—	+0.663220	−9.846544	−9.183324	22 506	568.630
10.0	4	4	+0.659302	−9.847228	−9.187926	25 288	719.068
10.0	4	8	+0.659456	−9.846967	−9.187511	20 766	785.202
10.0	8	—	+0.659267	−9.846237	−9.186970	34 537	610.037
10.0	8	4	+0.659308	−9.847178	−9.187871	23 178	762.927
10.0	8	8	+0.659146	−9.847090	−9.187944	29 346	822.300
10.0	8	16	+0.659161	−9.847202	−9.188041	23 407	978.041
10.0	16	—	+0.659727	−9.846250	−9.186523	22 091	735.663
10.0	16	8	+0.659073	−9.846888	−9.187815	18 146	955.435
10.0	16	16	+0.659103	−9.847266	−9.188162	17 500	1153
10.0	32	—	+0.659058	−9.846543	−9.187485	54 691	917.252
10.0	32	16	+0.659331	−9.846907	−9.187576	15 394	1396

where $\psi_{jk} = \psi(x_j, y_k)$, $x_j = -x_{\max} + hj$, $y_k = -y_{\max} + hk$, and $x_{\max} = y_{\max}$. Accordingly, the discretized Hamiltonian $\tilde{H}$ reads

$$\tilde{H} = -\frac{1}{2h^2}\tilde{T}_1 - \frac{1}{2h^2}\tilde{T}_2 + \tilde{V}_{\text{ext}}^1 + \tilde{V}_{\text{ext}}^2 + \tilde{V}_{\text{int}}, \tag{16}$$
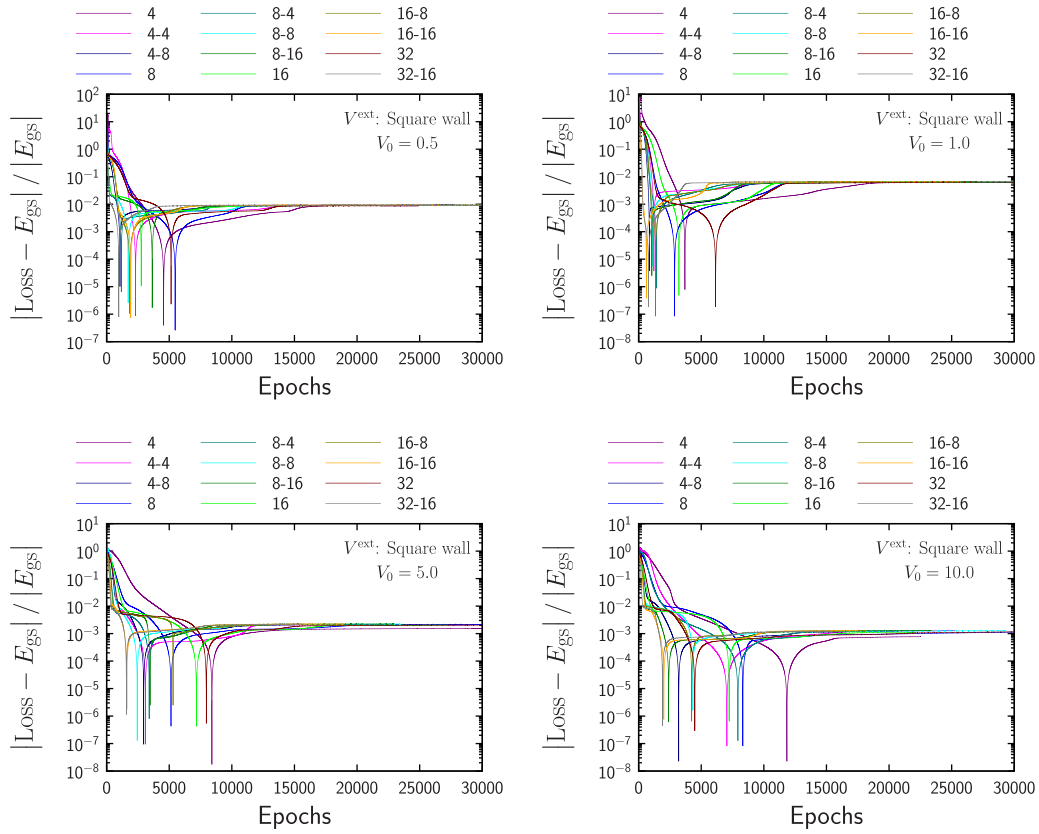
FIG. 6. Relative error of $\langle H \rangle$ to the exact ground-state energy E_{gs} for the square-well potential as functions of the number of epochs.

where $\tilde{T}_1$ and $\tilde{T}_2$ are the kinetic energy matrices

$$\tilde{T}_1 = T \otimes I_{M-1}, \tag{17a}$$

$$\tilde{T}_2 = I_{M-1} \otimes T, \tag{17b}$$

$\tilde{V}_{\mathrm{ext}}^1$ and $\tilde{V}_{\mathrm{ext}}^2$ are the external potential matrices

$$\tilde{V}_1^{\mathrm{ext}} = V^{\mathrm{ext}} \otimes I_{M-1}, \tag{18a}$$

$$\tilde{V}_2^{\mathrm{ext}} = I_{M-1} \otimes V^{\mathrm{ext}}, \tag{18b}$$

and $\tilde{V}_{\mathrm{int}}$ are the interaction matrix whose matrix elements are

$$(\tilde{V}_{\mathrm{int}})_{i+j(M-1),k+l(M-1)} = \begin{cases} \frac{1}{2}[V^{\mathrm{int}}(x_i, y_j) + V^{\mathrm{int}}(y_j, x_i)] = V^{\mathrm{int}}(x_i, y_j) & (\text{for } i = k, \ j = l), \\ 0 & (\text{otherwise}), \end{cases} \tag{19}$$

where I_{M-1} is the $(M-1) \times (M-1)$ identity matrix

$$I_{M-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{20}$$

and $\otimes$ is the Kronecker product. For instance, the matrix elements of Eqs. (17) read

$$(\tilde{T}_1)_{i+j(M-1),k+l(M-1)} = \begin{cases} -2 & (\text{for } i = k, \ j = l), \\ 1 & (\text{for } i = k, \ j = l \pm 1), \\ 0 & (\text{otherwise}), \end{cases} \tag{21a}$$

$$(\tilde{T}_2)_{i+j(M-1),k+l(M-1)} = \begin{cases} -2 & (\text{for } i = k, \ j = l), \\ 1 & (\text{for } i = k \pm 1, \ j = l), \\ 0 & (\text{otherwise}). \end{cases} \tag{21b}$$
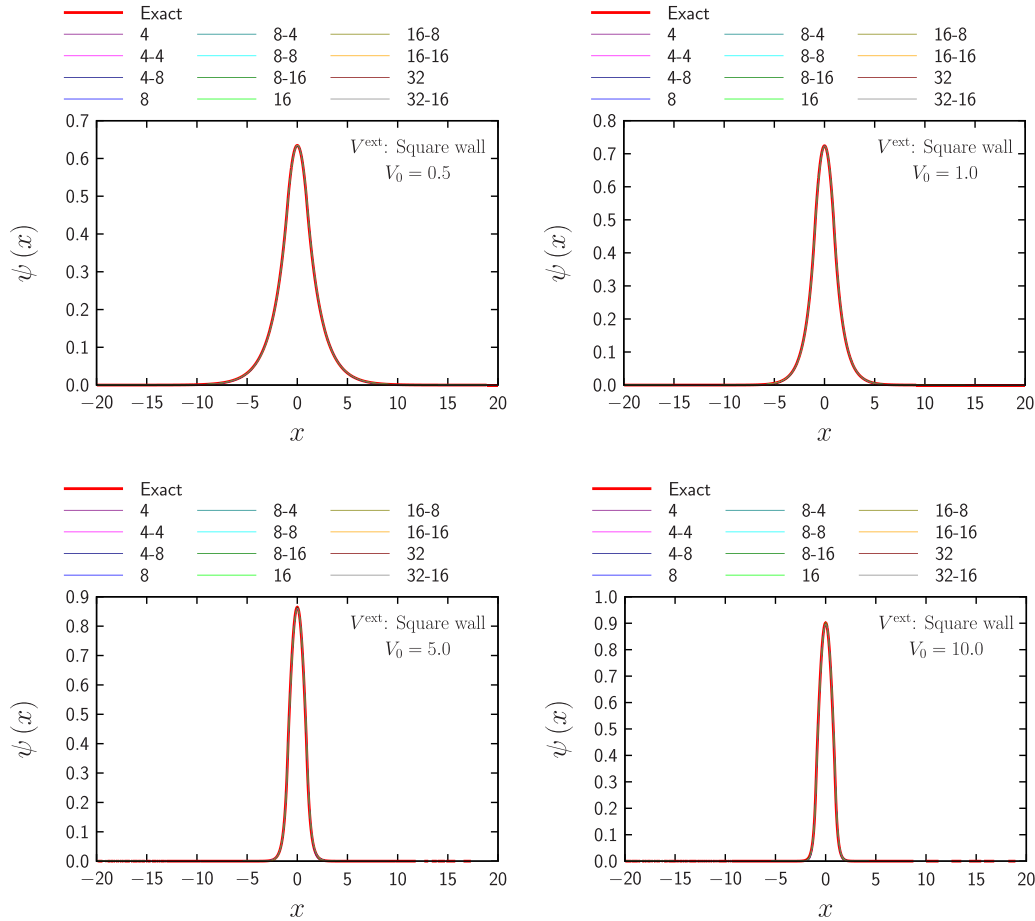
FIG. 7. Wave function under the square-well potential. The red thick line corresponds to the exact solution obtained by the exact diagonalization, while thin lines correspond to results of DNN calculation. Different thin line corresponds to different number of units. Almost all lines overlap with each other.

2. Symmetrization and antisymmetrization

The discretized wave function $\tilde{\psi}$ should be symmetrized or antisymmetrized. In general, for the arbitrary function $f(x, y)$, $f(x, y) + f(y, x)$ is a symmetrized function and $f(x, y) - f(y, x)$ is an antisymmetrized function.

To perform (anti)symmetrization in TENSORFLOW, instead of the simple $\tilde{\psi}$,

$$\tilde{\psi}_\pm = \begin{pmatrix} \psi_{11} \\ \psi_{12} \\ \vdots \\ \psi_{1(M-1)} \\ \psi_{21} \\ \psi_{22} \\ \vdots \\ \psi_{2(M-1)} \\ \vdots \\ \psi_{(M-1)1} \\ \psi_{(M-1)2} \\ \vdots \\ \psi_{(M-1)(M-1)} \end{pmatrix} \pm \begin{pmatrix} \psi_{11} \\ \psi_{21} \\ \vdots \\ \psi_{(M-1)1} \\ \psi_{12} \\ \psi_{22} \\ \vdots \\ \psi_{(M-1)2} \\ \vdots \\ \psi_{1(M-1)} \\ \psi_{2(M-1)} \\ \vdots \\ \psi_{(M-1)(M-1)} \end{pmatrix} \tag{22}$$

is assumed to be a trial wave function. In the TENSORFLOW code, instead of the original `predicts` and `output_wf`, $\tilde{\psi}_\pm$ is used in the third (the calculation process of the loss function) and fifth (calculate the ground-state energy) steps in Sec. II A. Note that this process can be easily done by using the following commands:

(1) `predicts_transpose = tf.reshape(predicts, [m, m])`,
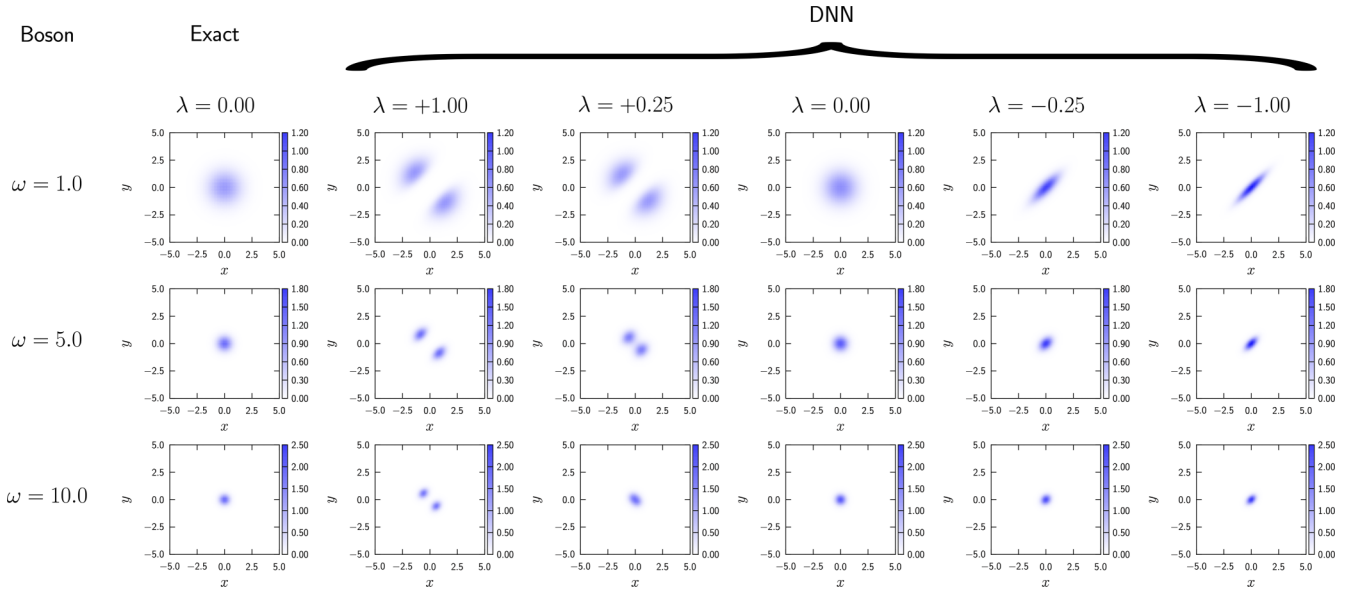(2) `predicts_transpose = tf.transpose(predicts_transpose)`,

FIG. 8. Two-body wave function under the harmonic oscillator potential for bosonic systems. The exact wave function without the interaction is shown in the left-most column.

(3) `predicts_transpose = tf.reshape(predicts_transpose, [m**2, 1])`,

(4) `predicts = tf.add(predicts, predicts_transpose)` for bosonic systems or `predicts = tf.subtract(predicts, predicts_transpose)` for fermionic systems,

(5) the final `predicts` is used to evaluate the loss function.

Here, m corresponds to the number of meshes M. Note that this method can be straightforwardly extended to multibody systems.

3. Two-body systems

Two-body systems under the harmonic oscillator potential [Eq. (8)] is tested. If there is no interaction $V^{int} \equiv 0$, the ground-state wave function ψ_{gs} and energy E_{gs} are known exactly. If one considers bosonic systems, they read

$$\psi_{gs}(x, y) = \sqrt{\frac{\omega}{\pi}} \exp\left[-\frac{\omega(x^2 + y^2)}{2}\right], \quad (23a)$$

$$E_{gs} = \omega, \quad (23b)$$

and if one considers the fermionic systems, they read

$$\psi_{gs}(x, y) = \frac{\omega}{\sqrt{\pi}}(x - y) \exp\left[-\frac{\omega(x^2 + y^2)}{2}\right], \quad (24a)$$

$$E_{gs} = 2\omega. \quad (24b)$$

Figures 8 and 9, respectively, show wave functions for bosonic and fermionic systems obtained in this work. The total energies and the calculation time are shown in Table III. For comparison, the exact wave functions [Eq. (23a) or Eq. (24a)] are also shown. Here, $x_{max} = y_{max} = 5$ and $M_x = M_y = 256$ are used for the spatial mesh and two layers each of which contains 32 units are used for the DNN. For the interaction, the Gaussian-type interaction

$$V^{int}(x, y) = \lambda \exp(-(x - y)^2) \quad (25)$$

is used, where λ is the strength of the interaction. The DNN results with $\lambda = 0$ show a good agreement with the exact results, demonstrating that the DNN technique works well.

Behavior of wave functions for nonzero λ is consistent qualitatively with our expectation: If λ is negative, i.e., the interaction is attractive, the wave function tends to collapse; if λ is positive, i.e., the interaction is repulsive, the wave function tends to be broad. Time cost per epoch is almost universal among all the calculation, while more epochs are required to reach convergence for fermionic systems than for bosonic systems. This may be because all the values are positive for the initial condition, while there are negative values for fermionic ground-state wave functions. More epochs are required for the repulsive interaction ($\lambda > 0$) than for the attractive interaction ($\lambda < 0$). This may be because the topology of the wave function is more complicated and extended in the repulsive case than the attractive case.

Finally, we point out a strange behavior of the obtained wave function of the two-body system of $\omega = 1.0$ without the interaction. Here, for simplicity, the two-layer DNN in which each layer is composed of four units is used. In the case of two layers, the function obtained by optimized weights of DNN is

$$u_{gs}(x, y) = A\,\text{softplus}\left(\sum_{j=1}^{N_{unit}} w_{2j}u_{2j}(x, y) + b_2\right), \quad (26a)$$

$$u_{2j}(x, y) = \text{softplus}\left(\sum_{k=1}^{N_{unit}} w_{1jk}u_{1k}(x, y) + b_{1j}\right), \quad (26b)$$

$$u_{1k}(x, y) = \text{softplus}(w_{0k0}x + w_{0k1}y + b_{0k}), \quad (26c)$$
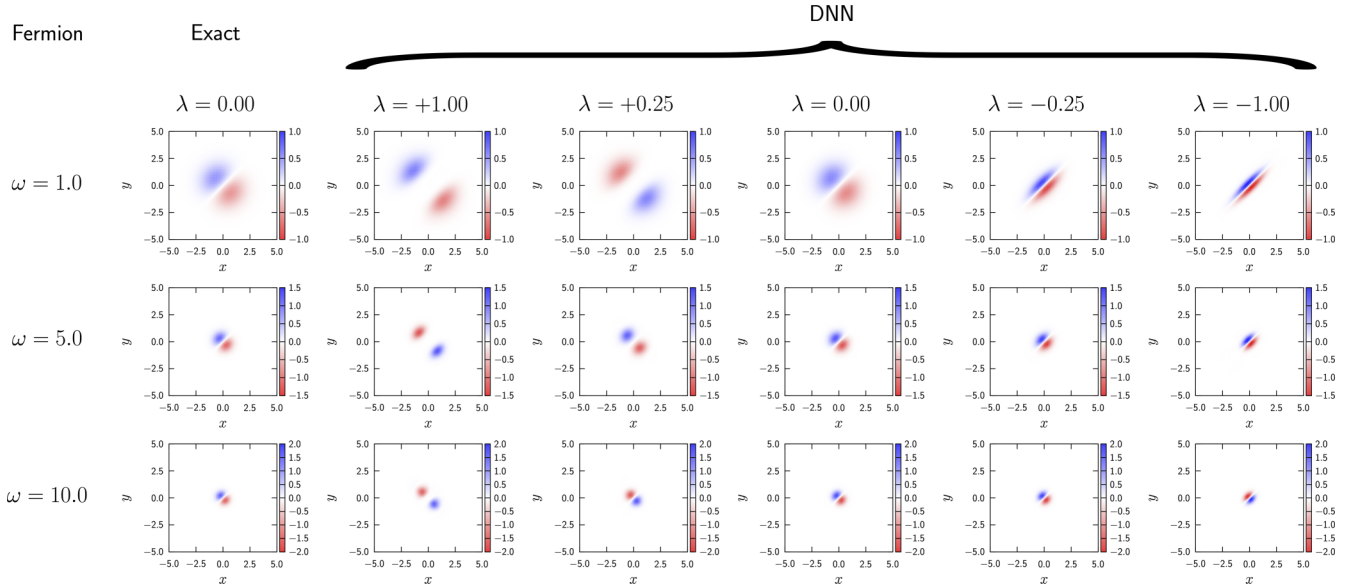
FIG. 9. Same as Fig. 8 but for fermionic systems.

TABLE III. Calculation summary of a two-body problem under the harmonic oscillator potential.

| Particles	ω	λ	Energy	# of Epochs	Time per Epoch (ms)
Boson	1.0	-1.00	-89.869381	8 490	23.583
Boson	1.0	-0.25	-19.848949	3 595	23.618
Boson	1.0	$+0.00$	$+0.999927$	27 242	23.424
Boson	1.0	$+0.25$	$+3.298725$	20 040	23.529
Boson	1.0	$+1.00$	$+3.835173$	21 763	23.677
Boson	5.0	-1.00	-87.554311	10 194	23.573
Boson	5.0	-0.25	-17.203647	10 893	23.712
Boson	5.0	$+0.00$	$+4.997829$	24 477	23.648
Boson	5.0	$+0.25$	$+21.149827$	20 721	23.797
Boson	5.0	$+1.00$	$+31.804917$	21 718	23.591
Boson	10.0	-1.00	-84.129658	18 635	23.566
Boson	10.0	-0.25	-13.118213	24 380	23.601
Boson	10.0	$+0.00$	$+9.991009$	27 489	23.692
Boson	10.0	$+0.25$	$+32.287424$	23 810	23.816
Boson	10.0	$+1.00$	$+72.688350$	19 591	23.601
Fermion	1.0	-1.00	-71.409493	18 794	23.928
Fermion	1.0	-0.25	-11.369632	17 999	23.818
Fermion	1.0	$+0.00$	$+1.999931$	19 215	23.843
Fermion	1.0	$+0.25$	$+3.298786$	25 187	23.804
Fermion	1.0	$+1.00$	$+3.839178$	106 163	24.136
Fermion	5.0	-1.00	-68.409494	18 558	23.731
Fermion	5.0	-0.25	-7.207718	19 618	23.787
Fermion	5.0	$+0.00$	$+9.995902$	75 208	23.841
Fermion	5.0	$+0.25$	$+21.385884$	51 691	23.607
Fermion	5.0	$+1.00$	$+31.804915$	27 885	23.799
Fermion	10.0	-1.00	-63.026667	26 603	23.894
Fermion	10.0	-0.25	$+0.302312$	21 783	23.805
Fermion	10.0	$+0.00$	$+19.975414$	10 135	23.741
Fermion	10.0	$+0.25$	$+38.060907$	79 973	24.054
Fermion	10.0	$+1.00$	$+73.245097$	59 289	23.972
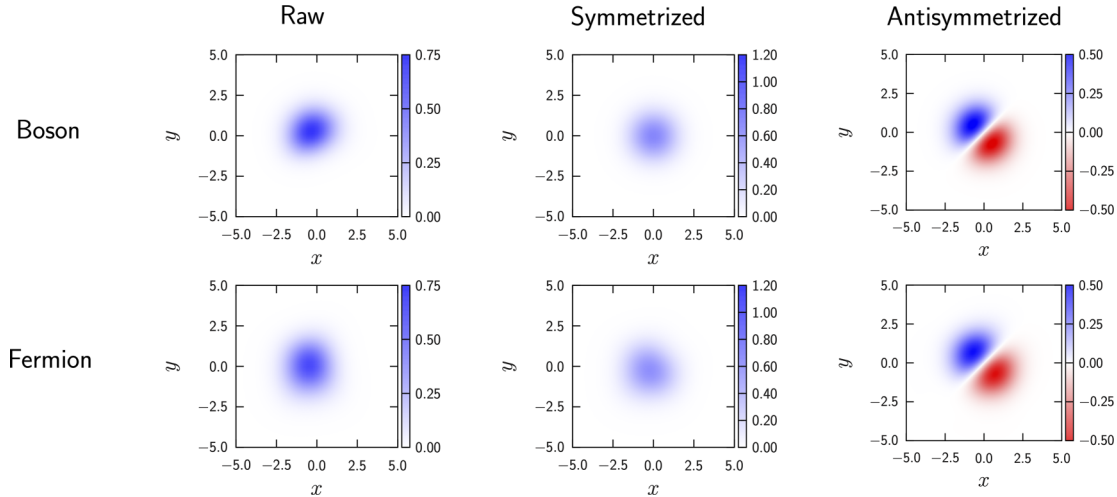
FIG. 10. DNN wave function of the raw [u_{gs} in Eq. (26a)], the symmetrized, and the antisymmetrized wave functions. The rows named "Boson" and "Fermion," respectively, correspond to the results obtained by minimizing the bosonic or fermionic energy expectation value.

where A is the normalization constant, N_{unit} is the number of units of each layer, w is a weight, and b is a bias. The left column of Fig. 10 shows $u_{gs}(x, y)$. The upper and lower rows, respectively, correspond to the results with minimizing the bosonic or fermionic energy expectation value. It is shown that the obtained function u_{gs}, which is referred to as the raw *wave function*, is not symmetric nor antisymmetric. After the symmetrization $\psi_{sym}(x, y) = [u_{gs}(x, y) + u_{gs}(y, x)]/A_{sym}$ or the antisymmetrization $\psi_{antisym}(x, y) = [u_{gs}(x, y) - u_{gs}(y, x)]/A_{antisym}$ is performed with the normalization constant A_{sym} or $A_{antisym}$, ψ_{sym} or $\psi_{antisym}$ can be regarded as the bosonic or fermionic ground-state wave function, respectively. The energy expectation value of the raw (u_{gs}), the symmetrized (ψ_{sym}), and the antisymmetrized ($\psi_{antisym}$) wave functions are shown in Table IV. A surprising fact is that even if the raw *wave function* is obtained by minimizing the bosonic (fermionic) expectation value, fermionic (bosonic) energy expectation value is close to the correct fermionic (bosonic) energy eigenvalue, and *vice versa*. The (anti)symmetrization corresponds to the projection of the raw *wave function* to the boson (fermion) subspace, while the remaining part is not supposed to be optimized well. This unexpected coincidence may be due to the smallness of the number of parameters in the DNN architecture, and deserves further study.

4. Three-body systems

Three-body systems under the harmonic oscillator potential [Eq. (8)] is tested. For simplicity, we consider a system

TABLE IV. Energy expectation value of the raw [u_{gs} in Eq. (26a)], the symmetrized, and the antisymmetrized wave functions. The rows named "Boson" and "Fermion," respectively, correspond to the results obtained by minimizing the bosonic or fermionic energy expectation value.

	Raw	Symmetrized	Antisymmetrized
Boson	1.13190	1.00012	2.05195
Fermion	1.13628	1.08235	2.00056

without any interaction $V^{int} \equiv 0$. Then, the ground-state wave function ψ_{gs} and energy E_{gs} are known exactly as

$$\psi_{gs}(x, y, z) = \left(\frac{\omega}{\pi}\right)^{3/4} \exp\left[-\frac{\omega(x^2 + y^2 + z^2)}{2}\right], \quad (27a)$$

$$E_{gs} = \frac{3}{2}\omega, \quad (27b)$$

for bosonic systems and

$$\psi_{gs}(x, y, z) = \left(\frac{\omega}{\pi}\right)^{3/4} \sqrt{\frac{\omega}{6}}[(x - y)(1 - 2\omega z^2)$$

$$+ (y - z)(1 - 2\omega x^2) + (z - x)(1 - 2\omega y^2)]$$

$$\times \exp\left[-\frac{\omega(x^2 + y^2 + z^2)}{2}\right], \quad (28a)$$

$$E_{gs} = \frac{9}{2}\omega \quad (28b)$$

for fermionic systems.

Figures 11 and 12, respectively, show wave functions for bosonic and fermionic systems obtained by this work. The total energies and the calculation time are shown in Table V. Here, $x_{max} = y_{max} = z_{max} = 5$ and $M_x = M_y = M_z = 64$ are used for the spatial mesh and two layers each of
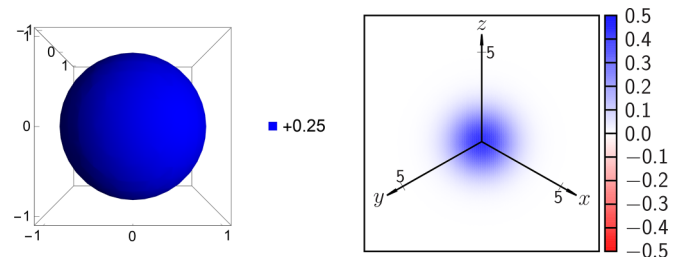

FIG. 11. (Left) Three-body wave function under the harmonic oscillator potential without interparticle interaction for bosonic systems. (Right) Slice of the three-body wave function at the plane $x + y + z = 0$.
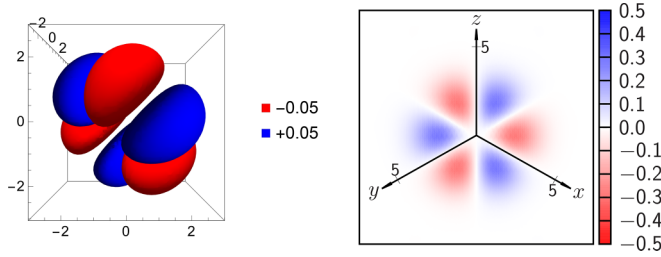
FIG. 12. Same as Fig. 11 but for fermionic systems.

which contains 32 units are used for the DNN. The interaction is not considered.

The DNN calculations reproduce the exact ground-state energies. The DNN wave functions are consistent with the exact solution. The number of epochs for three-body systems is comparable with those for two-body systems, where the number of units and layers are identical for these two cases. In contrast, the time per epoch for the three-body systems are about four times of that for the two-body systems. This is related to the number of spatial meshes: $256 \times 256 = 65\,536$ meshes are used for the two-body systems and $64 \times 64 \times 64 = 262\,144$ meshes are used for the three-body systems; thus, the number of meshes for the three-body systems are four times more than those for the two-body systems. Hence, it can be concluded that the time per epoch is almost proportional to the number of spatial meshes. This is reasonable since we use numerical methods for sparse matrices, in which the number of the nonzero matrix elements is $O(M^{Nd})$.

III. EXCITED-STATE CALCULATION

In this section, based on the variational principle, a method to calculate low-lying excited states sequentially is explained. Assume that wave functions of the ground state and n excited states, $|\psi_0\rangle$, $|\psi_1\rangle$, ..., $|\psi_n\rangle$, are obtained, where $|\psi_0\rangle = |\psi_{gs}\rangle$. We consider a problem of finding the $(n+1)$th excited state $|\psi_{n+1}\rangle$, which satisfies the orthnormal condition

$$\langle \psi_j | \psi_{n+1} \rangle = \delta_{j,n+1}, \tag{29}$$

by using a trial wave function $|\psi\rangle$. The $(n+1)$th wave function can be obtained with minimizing the expectation value

$$\langle H \rangle = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}, \tag{30}$$

where $|\psi\rangle$ is assumed to be orthogonal to $|\psi_j\rangle$ ($j = 0, 1, ..., n$). This can be implemented in TENSORFLOW with

TABLE V. Calculation summary of a three-body problem under the harmonic oscillator potential. Calculation is performed with $\omega = 1.0$.

Particles	Energy	# of Epochs	Time per Epoch (ms)
Boson	+1.497880	20 183	101.216
Fermion	+4.486830	22 770	98.356

assuming that

$$|\psi\rangle - \sum_{j=0}^{n} \langle \psi_j | \psi \rangle |\psi_j\rangle \tag{31}$$

is a trial wave function, instead of the simple $|\psi\rangle$. For one-body problem, $x_{\max} = 5$ and $M = 1024$ are used for the spatial mesh and the single-layer DNN with eight units is adopted.

A. Harmonic oscillator potential

One-body one-dimensional harmonic oscillators are taken as examples. The exact wave functions for several low-lying excited states are [41]

$$\psi_0(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \exp\left(-\frac{\omega x^2}{2}\right), \tag{32a}$$

$$\psi_1(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \sqrt{2\omega} x \exp\left(-\frac{\omega x^2}{2}\right), \tag{32b}$$

$$\psi_2(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \frac{2\omega x^2 - 1}{\sqrt{2}} \exp\left(-\frac{\omega x^2}{2}\right), \tag{32c}$$

$$\psi_3(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \sqrt{\frac{\omega}{3}} (2\omega x^2 - 3)x \exp\left(-\frac{\omega x^2}{2}\right), \tag{32d}$$

$$\psi_4(x) = \left(\frac{\omega}{\pi}\right)^{1/4} \frac{4\omega^2 x^4 - 12\omega x^2 + 3}{2\sqrt{6}} \exp\left(-\frac{\omega x^2}{2}\right), \tag{32e}$$

where ψ_n is the nth excited state, and the energies are

$$E_n = \left(n + \tfrac{1}{2}\right)\omega. \tag{33}$$

Figure 13 shows the wave functions of the ground state and first, second, third, and fourth excited states. Table VI shows the summary of calculations. Basically, not only the ground-state but also low-lying excited-states wave functions and energies are successfully calculated. Thus, it can be concluded that the method to calculate low-lying excited states proposed here works well.

The number of epochs are almost universal for all states calculated here. In contrast, the time per epoch for a higher excited state is slightly longer since calculation for orthogonal condition [Eq. (31)] is needed to be performed, while it takes just a few μs.

Let us explain why our simple DNN can describe even the excited states correctly. For simplicity, the single-layer DNN with the four unit is used. The optimized raw wave functions for the ground state (u_{gs}) and the first (u_{1st}) and the second (u_{2nd}) excited states are shown in Fig. 14. The optimized raw wave function for the nth excited state is

$$u_{n\text{th}}(x) = \sum_{j=0}^{n} a_j \psi_{n\text{th}}(x), \tag{34}$$

with $\sum_{j=0}^{n} |a_j|^2 = 1$, where $\psi_{n\text{th}}$ is the nth excited-state wave function. In the case of the first and second excited states,

$$u_{1\text{st}}(x) = 0.981758\psi_{0\text{th}}(x) + 0.190136\psi_{1\text{st}}(x), \tag{35a}$$

$$u_{2\text{nd}}(x) = 0.969707\psi_{0\text{th}}(x) - 0.0445505\psi_{1\text{st}}(x)$$
$$+ 0.240173\psi_{2\text{nd}}(x) \tag{35b}$$

FIG. 14. Optimized raw wave function u obtained by the DNN for the ground-state (0th) and the first and the second excited states.
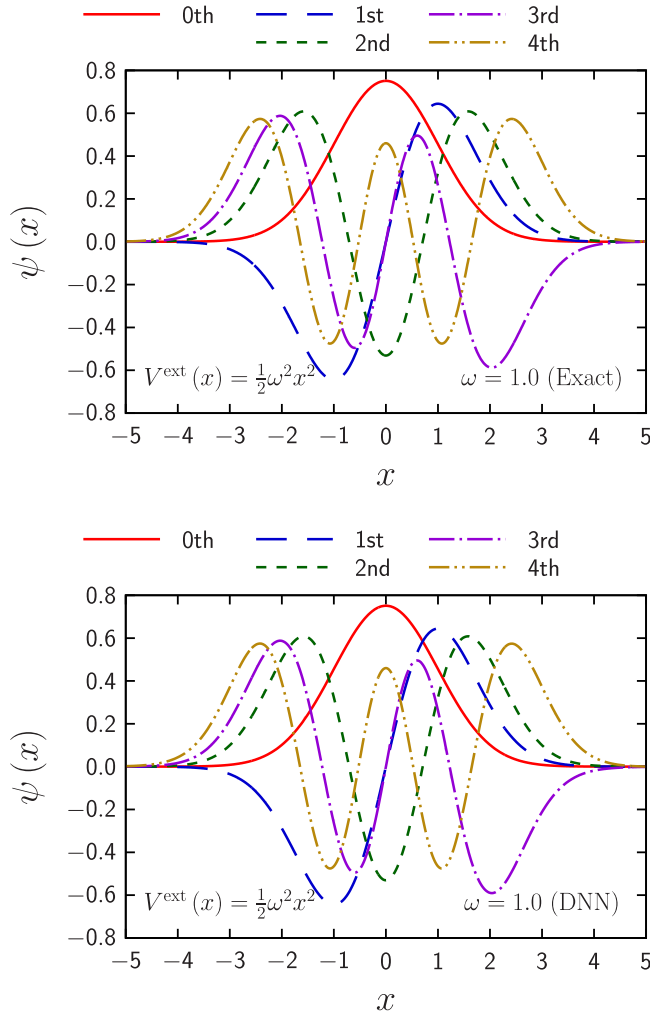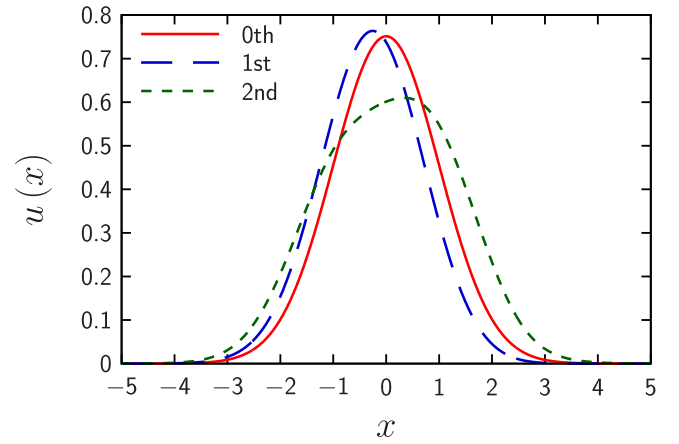
FIG. 13. Wave functions of the ground and low-lying excited states under the harmonic oscillator potential. The top panel shows the exact wave functions and the bottom one shows the DNN wave functions. To make consistency for the phase factor, $-\psi_3(x)$ is plotted for the exact wave function of the third excited state.

are obtained. We notice that the most part of the obtained raw *wave function* is the ground state and the small fraction is for the excited components; hence, the fairly simple raw *wave function*, which is made by the small architecture of the DNN and is close to that of the ground-state wave function, is capable of describing even the excited states.

TABLE VI. Calculation summary of excited states for a one-body problem under the harmonic oscillator potential. Calculation is performed with $\omega = 1.0$.

State	Energy	Epochs	Time per Epoch (μs)
0th	+0.499998	23 419	516.238
1st	+1.499991	25 646	519.026
2nd	+2.499986	23 157	527.849
3rd	+3.500193	37 880	534.115
4th	+4.500201	19 101	542.224

B. Double-well potential

To see the effect of degeneracy, we also test the double-well potential

$$V^{\text{ext}}(x) = (x^2 - \alpha^2)^2. \qquad (36)$$

If the central barrier is low enough, i.e., α is small enough, each state is not degenerate. In contrast, if the central barrier is high, i.e., α is large, low-lying excited states below the central barrier are twofold degenerate: One state ψ_{L} is localized into the left ($x < 0$) region while the other state ψ_{R} is localized into the right ($x > 0$) region, and $\psi_{\text{L}}(x) = \psi_{\text{R}}(-x)$ holds. Using a linear combination of these two degenerate states, one can recognize each state is degenerate of the following two states: $\psi_{\pm}(x) = [\psi_{\text{L}}(x) \pm \psi_{\text{R}}(x)]/\sqrt{2}$, where ψ_{+} (ψ_{-}) is a positive (negative) parity state. According to the exact diagonalization, $\alpha = 1.0$ and 1.25 give nondegenerate ground and low-lying excited states and thus ψ_j is just a jth excited state, while $\alpha = 2.0$ and 3.0 give ground and low-lying excited states, which are almost twofold degenerate: ψ_0 and ψ_1 correspond to the ground states and ψ_2 and ψ_3 correspond to the first excited states.

Figures 15 and 16 shows the wave functions of the ground state and first, second, third, and fourth excited states. Table VII shows the summary of calculations. Basically, not only the ground-state but also low-lying excited-states wave functions and energies are successfully calculated, even for the degenerate states. It is not apparent which calculation gives left-right bases (ψ_{L} and ψ_{R}), parity bases (ψ_{+} and ψ_{-}), or even general linear combinations. The DNN calculations for both $\alpha = 2.0$ and 3.0 obtained wave functions with the left-right bases, while it may depend on the initial condition. Note that the exact diagonalization for $\alpha = 2.0$ obtained wave functions with the parity bases, while wave functions with the left-right bases are plotted by using linear combinations in Fig. 16 to make a comparison with the DNN result easily.

C. Two-body systems

Two-body one-dimensional harmonic oscillators are taken as the last examples. Here, $x_{\max} = y_{\max} = 5$ and $M_x = M_y = 256$ are used for the spatial mesh and two layers each of which
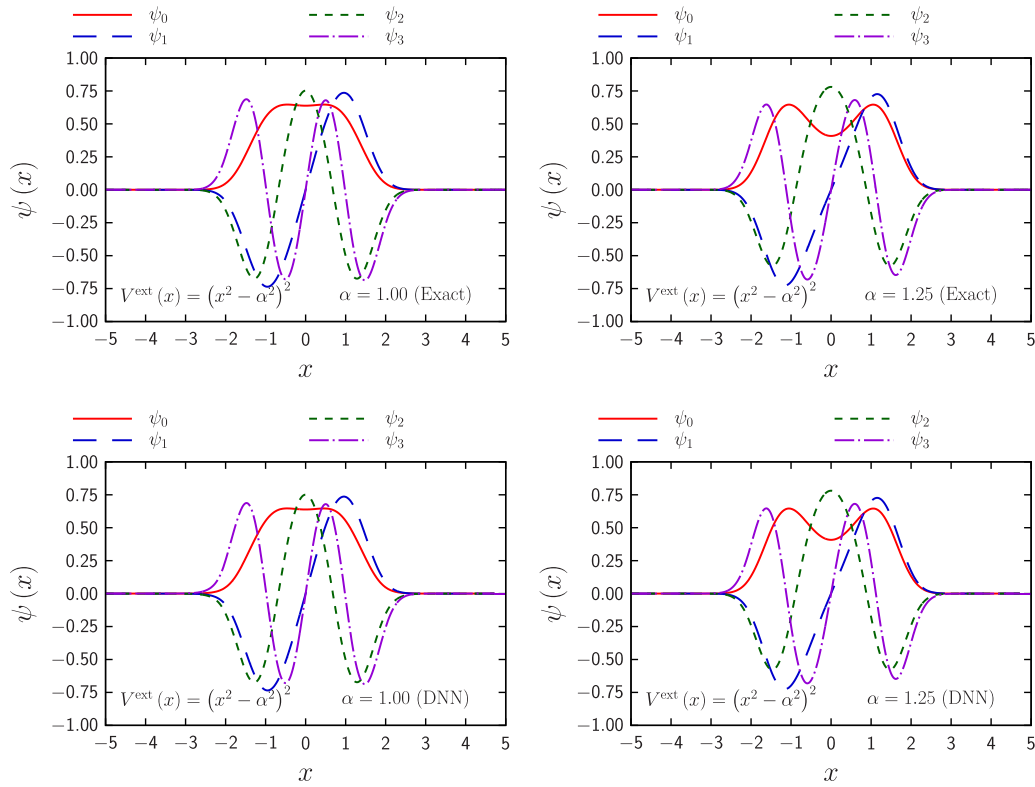
FIG. 15. Wave functions of the ground and low-lying excited states under the double-well potential for $\alpha = 1.0$ and 1.25. The top panels show the exact wave functions and the bottom ones show the DNN wave functions.
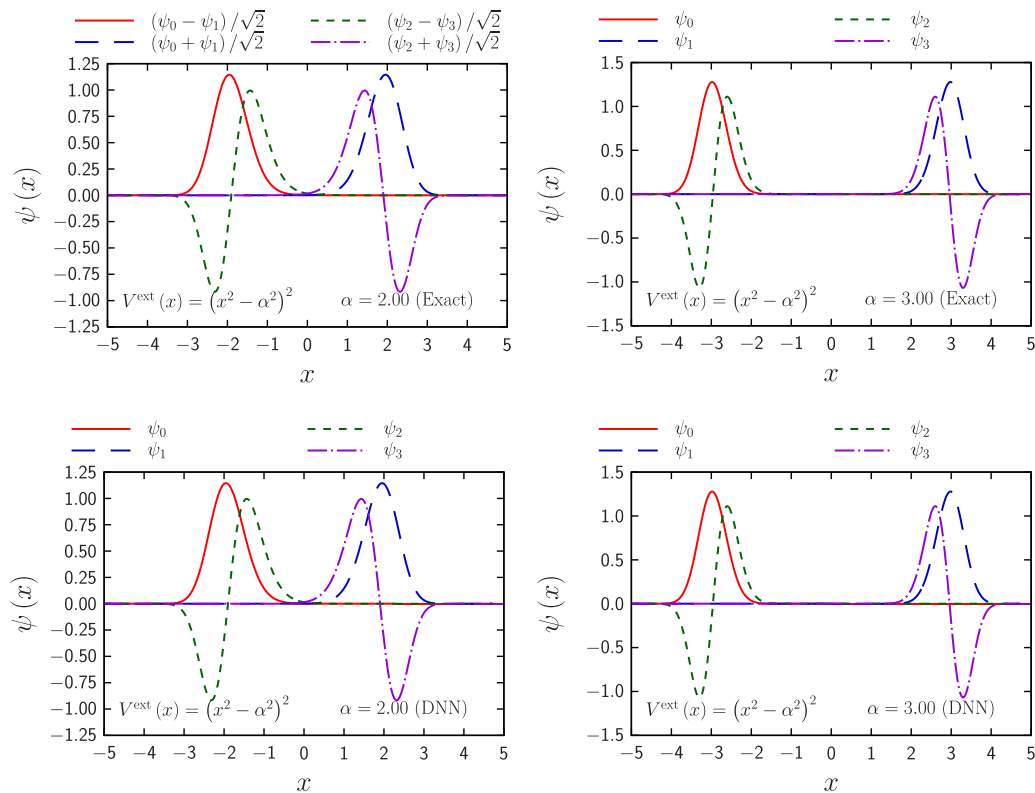

FIG. 16. Same as Fig. 15 but for $\alpha = 2.0$ and 3.0.

TABLE VII. Calculation summary of a one-body problem under the double-well potential.

		Energy			
α	j	Exact diagonalization	Deep neural network	Epochs	Time per Epoch (μs)
1.0	0	+0.869573	+0.869706	29 108	513.203
1.0	1	+1.661393	+1.661685	35 596	523.556
1.0	2	+3.543667	+3.544327	55 037	524.351
1.0	3	+5.665058	+5.666010	20 083	536.858
1.25	0	+1.417858	+1.417886	30 284	512.083
1.25	1	+1.725904	+1.726013	36 992	523.092
1.25	2	+3.717933	+3.717949	30 943	528.675
1.25	3	+5.424725	+5.424850	24 474	534.907
2.0	0	+2.762317	+2.762333	23 663	519.261
2.0	1	+2.762333	+2.762343	24 764	529.129
2.0	2	+7.988520	+7.989654	19 491	532.193
2.0	3	+7.990618	+7.989601	18 662	534.543
3.0	0	+4.214229	+4.214253	20 526	515.042
3.0	1	+4.214229	+4.214284	20 290	521.291
3.0	2	+12.526202	+12.526214	51 131	519.062
3.0	3	+12.526202	+12.526278	12 488	534.752

contains 32 units are used for the DNN. Here, the interparticle interaction is not considered. The exact wave functions for several low-lying excited states can be written as linear combinations of Eqs. (32):

$$\Psi_0(x, y) = \psi_0(x)\psi_0(y), \tag{37a}$$

$$\Psi_1(x, y) = \frac{1}{\sqrt{2}}[\psi_0(x)\psi_1(y) + \psi_1(x)\psi_0(y)], \tag{37b}$$

$$\Psi_2(x, y) = \frac{1}{\sqrt{2}}[\psi_0(x)\psi_2(y) + \psi_2(x)\psi_0(y)], \tag{37c}$$

$$\Psi_3(x, y) = \psi_1(x)\psi_1(y), \tag{37d}$$

where the energy eigenvalue of Ψ_0, Ψ_1, Ψ_2, and Ψ_3 are equal to, respectively, 1, 2, 3, and 3 for bosonic systems and

$$\Psi_0(x, y) = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_0(x) & \psi_1(x) \\ \psi_0(y) & \psi_1(y) \end{vmatrix}, \tag{38a}$$

$$\Psi_1(x, y) = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_0(x) & \psi_2(x) \\ \psi_0(y) & \psi_2(y) \end{vmatrix}, \tag{38b}$$

$$\Psi_2(x, y) = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_0(x) & \psi_3(x) \\ \psi_0(y) & \psi_3(y) \end{vmatrix}, \tag{38c}$$

$$\Psi_3(x, y) = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_1(x) & \psi_2(x) \\ \psi_1(y) & \psi_2(y) \end{vmatrix}, \tag{38d}$$
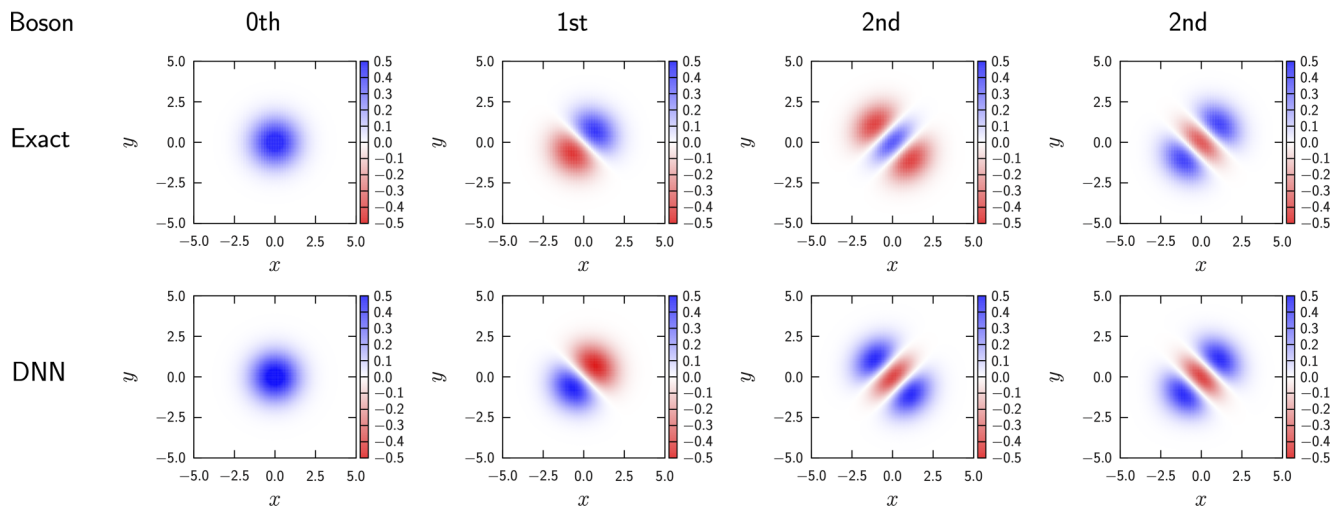

FIG. 17. Two-body wave function for the ground and low-lying excited states under the harmonic oscillator potential without the interaction for bosonic systems. The exact wave function is shown in the top row.
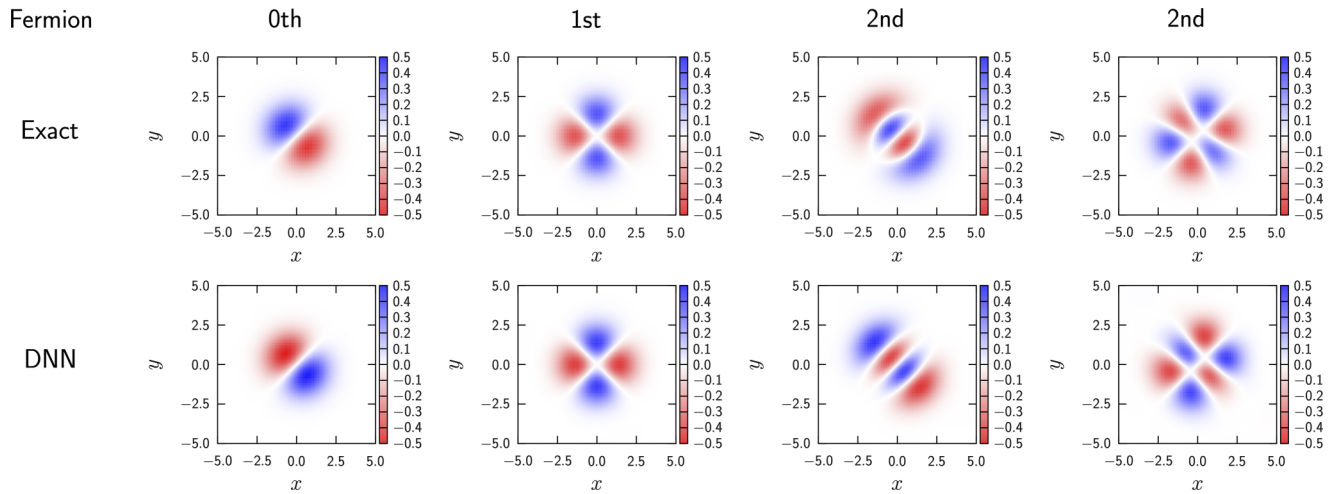
FIG. 18. Same as Fig. 17 but for fermionic systems.

where the energy eigenvalue of Ψ_0, Ψ_1, Ψ_2, and Ψ_3 are equal to, respectively, 2, 3, 4, and 4 for fermionic systems. Note that the second excited states, Ψ_2 and Ψ_3, are twofold degenerate in both the bosonic and fermionic systems.

Figures 17 and 18, respectively, show the wave functions of the ground state and first and second excited states. Table VIII shows the summary of calculations. Note that $[\Psi_2(x, y) \pm \Psi_3(x, y)]/\sqrt{2}$ are plotted for the second excited states for exact solutions. Not only the ground-state but also low-lying excited-states wave functions and energies are successfully calculated even for two-body systems. In addition, as one-body problems, the numerical cost for a low-lying excited state is almost the same as that for the ground state. Thus, this method to calculate low-lying excited states can work even for multibody systems with a reasonable numerical cost.

IV. SUMMARY

In this paper, we proposed a method to calculate the wave functions and energies of not only the ground state but also low-lying excited states of quantum multibody systems using the deep neural network and the unsupervised machine learning technique. To calculate systems of many-particle systems of identical particles, a simple method of symmetrization for bosonic systems and antisymmetrization for fermionic systems were also proposed.

The obtained wave functions and energies are consistent with the exact solution. We found that the neural network is not necessarily large for one-body systems, which also enables us to analyze the internal structure of the deep neural network used. For instance, just only one hidden layer with four units is enough to describe the ground-state wave function of the harmonic oscillator. This can be understood by using the piecewise approximation with linear functions. We confirmed that our simple (anti)symmetrization method works for multibody systems. The numerical cost per epoch for fermionic systems is almost the same as that for bosonic systems. The numerical cost is almost proportional to the number of spatial meshes since the sparse matrix representation is used. In addition, the numerical cost for a low-lying excited state is almost the same as that for the ground state.

The deep neural network has been applied to solve many-fermion systems where the ground-state wave function is assumed to be a Jastrow wave function [46,47,57]. The method proposed in this paper can be an alternative method to solve many-fermionic systems since the ansatz for the ground-state wave function is lenient, and the symmetrization and antisymmetrization are treated on an equal footing.

Since the numerical cost is not so large and our (anti)symmetrization is quite simple, this method can be an alternative method to calculate wave functions and energies of the ground and low-lying excited states, for instance, for the electronic structure of molecules and solids, for the nuclear structure of atomic nuclei including a tetra neutron [66–68], and for cold atoms [69].

At this moment, we only considered one-dimensional systems, while most problems interested are three-dimensional systems. In addition, spin components, or even isospin components for nuclear systems, are often important. The restricted Boltzmann machine has been applied to obtain the ground- and low-lying excited-states wave functions [45,70,71]. Since the input is discrete variables in the spin systems, the Boltzmann machine is suitable. Such pioneering works may help to consider the spin (or isospin) components

TABLE VIII. Calculation summary of excited states for a two-body problem under the harmonic oscillator potential. Calculation is performed with $\omega = 1.0$.

Particles	State	Energy	Epochs	Time per Epoch (ms)
Boson	0th	+0.999935	26 197	23.859
Boson	1st	+1.999778	26 224	24.228
Boson	2nd (1)	+2.999801	19 204	24.502
Boson	2nd (2)	+3.000229	15 239	25.450
Fermion	0th	+1.999855	38 500	23.844
Fermion	1st	+2.999771	29 552	24.205
Fermion	2nd (1)	+3.999251	28 197	24.674
Fermion	2nd (2)	+4.004341	11 871	25.772

in this work. Such extensions are possible within our framework, and remain for future work.

As far as we know, all the calculations using the deep neural network for wave functions are static, while describing many phenomena including the interaction between matter and laser [72,73], ion-cluster collision [74], heavy-ion collision [75], nuclear fission [76,77], and fusion [78]. To describe such phenomena, time evolution from a state obtained by the deep neural network is also interesting, while it is left for a future study.

Finally, let us make a comment on the interpretation of the wave functions obtained in our work. As we have shown, thanks to the simplicity of the deep neural network, we could interpret the structure of the network easily. We found that replacing the softmax function with the ReLU activation provides a piecewise linear function which approximates the ground-state wave function. Since any wave function including those for excited states, which is naturally continuous, can be approximated by a piecewise-linear function, we intuitively conclude that the neural network representation can work for any physical quantum mechanical system in any dimensions. The physical meaning of the piecewise-linear functions is as follows. First of all, linear functions are solutions of the free Schrödinger equation with no potential term. So it is a good idea to start with linear functions in physical systems. Then the inclusion of the potential term in the Hamiltonian causes the curvature of the wave function. The curvature is determined by the interplay between the Laplacian and the potential term in the Hamiltonian. So, the kink structure of the wave function is dictated by the Hamiltonian. The kinks correspond to the ReLU activations, thus in effect, the nonlinearity in the Hamiltonian corresponds to the neural network structure. This reminds us of the work [79] in which the deep layers of the deep Boltzmann machine representing the ground-state wave functions of spin systems were interpreted as a Euclidean Hamiltonian evolution, or the work [80–82] in which the deep layers of the sparse neural network used for the AdS/CFT correspondence were interpreted as a bulk curved geometry. Further interplay between the sparsity of the interpretable neural networks and Hamiltonians of physical systems is to be discovered.

ACKNOWLEDGMENTS

[1] L. D. Faddeev, Scattering theory for a three-particle system, Sov. Phys. JETP **12**, 1014 (1961) [Zh. Eksp. Teor. Fiz. **39**, 1459 (1961)].

[2] O. A. Yakubovskĭ, On the integral equations in the theory of *N* particle scattering, Sov. J. Nucl. Phys. **5**, 937 (1967) [J. Nucl. Phys. (U.S.S.R.) **5**, 1312 (1967)].

[3] M. Fabre de la Ripelle, The potential harmonic expansion method, Ann. Phys. **147**, 281 (1983).

[4] J. Carlson, Green's function Monte Carlo study of light nuclei, Phys. Rev. C **36**, 2026 (1987).

[5] M. Kamimura, Nonadiabatic coupled-rearrangement-channel approach to muonic molecules, Phys. Rev. A **38**, 621 (1988).

[6] H. Kamada and W. Glöckle, Solutions of the Yakubovsky equations for four-body model systems, Nucl. Phys. A **548**, 205 (1992).

[7] K. Varga and Y. Suzuki, Precise solution of few-body problems with the stochastic variational method on a correlated Gaussian basis, Phys. Rev. C **52**, 2885 (1995).

[8] M. Viviani, A. Kievsky, and S. Rosati, Calculation of the α-particle ground state, Few-Body Syst. **18**, 25 (1995).

[9] P. Navrátil and B. R. Barrett, Four-nucleon shell-model calculations in a Faddeev-like approach, Phys. Rev. C **59**, 1906 (1999).

[10] P. Navrátil, G. P. Kamuntavičius, and B. R. Barrett, Few-nucleon systems in a translationally invariant harmonic oscillator basis, Phys. Rev. C **61**, 044001 (2000).

[11] N. Barnea, W. Leidemann, and G. Orlandini, State dependent effective interaction for the hyperspherical formalism, Phys. Rev. C **61**, 054001 (2000).

[12] H. Kamada, A. Nogga, W. Glöckle, E. Hiyama, M. Kamimura, K. Varga, Y. Suzuki, M. Viviani, A. Kievsky, S. Rosati, J. Carlson, S. C. Pieper, R. B. Wiringa, P. Navrátil, B. R. Barrett, N. Barnea, W. Leidemann, and G. Orlandini, Benchmark test calculation of a four-nucleon bound state, Phys. Rev. C **64**, 044001 (2001).

[13] D. Ceperley, Ground state of the fermion one-component plasma: A Monte Carlo study in two and three dimensions, Phys. Rev. B **18**, 3126 (1978).

[14] D. M. Ceperley and B. J. Alder, Ground State of the Electron Gas by a Stochastic Method, Phys. Rev. Lett. **45**, 566 (1980).

[15] J. J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, Full configuration interaction perspective on the homogeneous electron gas, Phys. Rev. B **85**, 081103(R) (2012).

[16] J. J. Shepherd, G. H. Booth, and A. Alavi, Investigation of the full configuration interaction quantum Monte Carlo method using homogeneous electron gas models, J. Chem. Phys. **136**, 244101 (2012).

[17] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, Towards an exact description of electronic wave functions in real solids, Nature (London) **493**, 365 (2013).

[18] D. Lonardoni, S. Gandolfi, J. E. Lynn, C. Petrie, J. Carlson, K. E. Schmidt, and A. Schwenk, Auxiliary field diffusion Monte

Carlo calculations of light and medium-mass nuclei with local chiral interactions, Phys. Rev. C **97**, 044318 (2018).

[19] T. Shen, Y. Liu, Y. Yu, and B. M. Rubenstein, Finite temperature auxiliary field quantum Monte Carlo in the canonical ensemble, J. Chem. Phys. **153**, 204108 (2020).

[20] J. A. Pople, J. S. Binkley, and R. Seeger, Theoretical models incorporating electron correlation, Int. J. Quantum Chem. **10**, 1 (1976).

[21] J. A. Pople, R. Seeger, and K. Raghavachari, Variational configuration interaction methods and comparison with perturbation theory, Int. J. Quantum Chem. **12**, 149 (1977).

[22] J. A. Pople, Nobel Lecture: Quantum chemical models, Rev. Mod. Phys. **71**, 1267 (1999).

[23] F. Coester, Bound states of a many-particle system, Nucl. Phys. **7**, 421 (1958).

[24] F. Coester and H. Kümmel, Short-range correlations in nuclear wave functions, Nucl. Phys. **17**, 477 (1960).

[25] J. Čížek, On the correlation problem in atomic and molecular systems: Calculation of wave-function components in Ursell-type expansion using quantum-field theoretical methods, J. Chem. Phys. **45**, 4256 (1966).

[26] J. Čížek and J. Paldus, Correlation problems in atomic and molecular systems III. Rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methodst, Int. J. Quantum Chem. **5**, 359 (1971).

[27] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

[28] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[29] W. Kohn, Nobel Lecture: Electronic structure of matter—Wave functions and density functionals, Rev. Mod. Phys. **71**, 1253 (1999).

[30] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[31] S. R. White and R. L. Martin, *Ab initio* quantum chemistry using the density matrix renormalization group, J. Chem. Phys. **110**, 4127 (1999).

[32] U. Schollwöck, The density-matrix renormalization group, Rev. Mod. Phys. **77**, 259 (2005).

[33] A. Baiardi and M. Reiher, The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges, J. Chem. Phys. **152**, 040903 (2020).

[34] V. I. Anisimov, A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar, First-principles calculations of the electronic structure and spectra of strongly correlated systems: Dynamical mean-field theory, J. Phys.: Condens. Matter **9**, 7359 (1997).

[35] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Electronic structure calculations with dynamical mean-field theory, Rev. Mod. Phys. **78**, 865 (2006).

[36] G. Knizia and G. K.-L. Chan, Density Matrix Embedding: A Simple Alternative to Dynamical Mean-Field Theory, Phys. Rev. Lett. **109**, 186404 (2012).

[37] R. Brockmann and J. Frank, Lattice Quantum Hadrodynamics, Phys. Rev. Lett. **68**, 1830 (1992).

[38] J.-W. Chen and D. B. Kaplan, Lattice Theory for Low-Energy Fermions at Nonzero Chemical Potential, Phys. Rev. Lett. **92**, 257002 (2004).

[39] D. Lee, Lattice simulations for few- and many-body systems, Prog. Part. Nucl. Phys. **63**, 117 (2009).

[40] J. E. Drut and A. N. Nicholson, Lattice methods for strongly interacting many-body systems, J. Phys. G **40**, 043101 (2013).

[41] L. Schiff, *Quantum Mechanics*, 3rd ed. (McGraw-Hill, New York, NY, 1968).

[42] A. D. Becke, Perspective: Fifty years of density-functional theory in chemical physics, J. Chem. Phys. **140**, 18A301 (2014).

[43] R. Jastrow, Many-body problem with strong forces, Phys. Rev. **98**, 1479 (1955).

[44] P. Seth, P. L. Ríos, and R. J. Needs, Quantum Monte Carlo study of the first-row atoms and ions, J. Chem. Phys. **134**, 084105 (2011).

[45] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).

[46] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, *Ab initio* solution of the many-electron Schrödinger equation with deep neural networks, Phys. Rev. Res. **2**, 033429 (2020).

[47] J. Hermann, Z. Schätzle, and F. Noé, Deep-neural-network solution of the electronic Schrödinger equation, Nat. Chem. **12**, 891 (2020).

[48] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. **2**, 303 (1989).

[49] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks **4**, 251 (1991).

[50] T. Cheon, Green function Monte Carlo method for excited states of quantum system, Prog. Theor. Phys. **96**, 971 (1996).

[51] H. Masui and T. Sato, Monte Carlo study of bound states in a few-nucleon system: Method of continued fractions, Prog. Theor. Phys. **100**, 977 (1998).

[52] L. Hedin, On correlation effects in electron spectroscopies and the GW approximation, J. Phys.: Condens. Matter **11**, R489 (1999).

[53] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Self-consistent mean-field models for nuclear structure, Rev. Mod. Phys. **75**, 121 (2003).

[54] G. Colò, P. F. Bortignon, H. Sagawa, K. Moghrabi, M. Grasso, and N. Van Giai, Microscopic theory of particle-vibration coupling, J. Phys.: Conf. Ser. **321**, 012018 (2011).

[55] T. Nakatsukasa, K. Matsuyanagi, M. Matsuo, and K. Yabana, Time-dependent density-functional description of nuclear dynamics, Rev. Mod. Phys. **88**, 045004 (2016).

[56] L. Reining, The GW approximation: Content, successes and limitations, WIREs Comput. Mol. Sci. **8**, e1344 (2018).

[57] M. T. Entwistle, Z. Schätzle, P. A. Erdman, J. Hermann, and F. Noé, Electronic excited states in deep variational Monte Carlo, Nat. Commun. **14**, 274 (2023).

[58] H. Saito, Method to solve quantum few-body problems with artificial neural networks, J. Phys. Soc. Jpn. **87**, 074002 (2018).

[59] J. W. T. Keeble and A. Rios, Machine learning the deuteron, Phys. Lett. B **809**, 135743 (2020).

[60] Y. Wang, Y. Liao, and H. Xie, Solving Schrödinger equation using tensor neural network, arXiv:2209.12572 [physics.comp-ph].

[61] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg *et al.*, Tensor-

Flow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.

[62] D. P. Kingma and J. Ba, A method for stochastic optimization, in *Third International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA,May 7–9, 2015, Conference Track Proceedings*, edited by Y. Bengio and Y. Le-Cun (2015), arXiv:1412.6980 [cs.LG].

[63] N. Tajima, Hartree-Fock+BCS Approach to Unstable Nuclei with the Skyrme Force, Prog. Theor. Phys. Suppl. **142**, 265 (2001).

[64] If $\tilde{H}$ is directly discretized, wave functions of the ground state and $(M - 2)$ excited states can be obtained.

[65] Due to the internal code of the TENSORFLOW, when the error of the loss function is calculated, a single precision floating point number (float32) seems to be used. Thus, the error of the loss function becomes zero if its actual value is smaller than about 1.0×10^{-7}. Note that the convergence criteria of the DFT calculation is often about 1.0×10^{-8} or even 1.0×10^{-10}; for instance, see the sample input of Ref. [83].

[66] K. Kisamori, S. Shimoura, H. Miya, S. Michimasa, S. Ota, M. Assie, H. Baba, T. Baba, D. Beaumel, M. Dozono, T. Fujii, N. Fukuda, S. Go, F. Hammache, E. Ideguchi, N. Inabe, M. Itoh, D. Kameda, S. Kawase, T. Kawabata *et al.*, Candidate Resonant Tetraneutron State Populated by the $^4\mathrm{He}(^8\mathrm{He},^8\mathrm{Be})$ Reaction, Phys. Rev. Lett. **116**, 052501 (2016).

[67] T. Faestermann, A. Bergmaier, R. Gernhäuser, D. Koll, and M. Mahgoub, Indications for a bound tetraneutron, Phys. Lett. B **824**, 136799 (2022).

[68] M. Duer, T. Aumann, R. Gernhäuser, V. Panin, S. Paschalis, D. M. Rossi, N. L. Achouri, D. Ahn, H. Baba, C. A. Bertulani, M. Böhmer, K. Boretzky, C. Caesar, N. Chiga, A. Corsi, D. Cortina-Gil, C. A. Douma, F. Dufter, Z. Elekes, J. Feng *et al.*, Observation of a correlated free four-neutron system, Nature (London) **606**, 678 (2022).

[69] A. N. Wenz, G. Zürn, S. Murmann, I. Brouzos, T. Lompe, and S. Jochim, From few to many: Observing the formation of a Fermi sea one atom at a time, Science **342**, 457 (2013).

[70] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Symmetries and Many-Body Excitations with Neural-Network Quantum States, Phys. Rev. Lett. **121**, 167204 (2018).

[71] Y. Nomura, Machine learning quantum states— Extensions to Fermion–Boson coupled systems and

excited-state calculations, J. Phys. Soc. Jpn. **89**, 054706 (2020).

[72] S. A. Sato, K. Yabana, Y. Shinohara, T. Otobe, K.-M. Lee, and G. F. Bertsch, Time-dependent density functional theory of high-intensity short-pulse laser irradiation on insulators, Phys. Rev. B **92**, 205413 (2015).

[73] Y. Tanaka and S. Tsuneyuki, Development of the temperature-dependent interatomic potential for molecular dynamics simulation of metal irradiated with an ultrashort pulse laser, J. Phys.: Condens. Matter **34**, 165901 (2022).

[74] K. Yabana, T. Tazawa, Y. Abe, and P. Bożek, Time-dependent mean-field description for multiple electron transfer in slow ion-cluster collisions, Phys. Rev. A **57**, R3165 (1998).

[75] K. Sekizawa and K. Yabana, Time-dependent Hartree-Fock calculations for multinucleon transfer processes in $^{40,48}\mathrm{Ca} + {}^{124}\mathrm{Sn}$, $^{40}\mathrm{Ca} + {}^{208}\mathrm{Pb}$, and $^{58}\mathrm{Ni} + {}^{208}\mathrm{Pb}$ reactions, Phys. Rev. C **88**, 014614 (2013).

[76] B. B. Back, H. Esbensen, C. L. Jiang, and K. E. Rehm, Recent developments in heavy-ion fusion reactions, Rev. Mod. Phys. **86**, 317 (2014).

[77] J. Zhao, J. Xiang, Z.-P. Li, T. Nikšić, D. Vretenar, and S.-G. Zhou, Time-dependent generator-coordinate-method study of mass-asymmetric fission of actinides, Phys. Rev. C **99**, 054613 (2019).

[78] K. Sekizawa and K. Hagino, Time-dependent Hartree-Fock plus Langevin approach for hot fusion reactions to synthesize the $Z = 120$ superheavy element, Phys. Rev. C **99**, 051602(R) (2019).

[79] G. Carleo, Y. Nomura, and M. Imada, Constructing exact representations of quantum many-body systems with deep neural networks, Nat. Commun. **9**, 5322 (2018).

[80] K. Hashimoto, S. Sugishita, A. Tanaka, and A. Tomiya, Deep learning and the AdS/CFT correspondence, Phys. Rev. D **98**, 046019 (2018).

[81] K. Hashimoto, S. Sugishita, A. Tanaka, and A. Tomiya, Deep learning and holographic QCD, Phys. Rev. D **98**, 106014 (2018).

[82] K. Hashimoto, AdS/CFT correspondence as a deep Boltzmann machine, Phys. Rev. D **99**, 106017 (2019).

[83] T. Ozaki, K. Hino, H. Kawai, and M. Toyoda, AD-PACK Ver. 2.2, https://www.openmx-square.org/adpack_man2.2/adpack2_2.html.