Neural-network quantum states for solving few-body problems: Application to Efimov physics

Sora Yokoi, Shimpei Endo,* and Hiroki Saito†
Department of Engineering Science, University of Electro-Communications, Tokyo 182-8585, Japan

(Received 5 April 2026; accepted 28 August 2026; published 18 September 2026)

Neural-network quantum states have been developed as an efficient method for solving quantum many-body problems, not only in lattice systems but also in systems of particles in continuous space. Here, we apply this approach to strongly interacting few-body problems in continuous space at unitarity: the Efimov states and associated few-body bound states. We extend the previous work [J. Phys. Soc. Jpn. **87**, 074002 (2018)], in which the ground states of few-boson systems were obtained, to the first excited states with a projection method, and also to a mass-imbalanced fermionic system consisting of two identical fermions and a third particle. The obtained energies of the ground and first excited states of these systems agree well with previously reported results. Furthermore, the proposed approach also reproduces the discrete scale invariance between the ground and first excited states and the critical-mass behavior in mass-imbalanced fermionic systems.

DOI: 10.1103/31mv-7tz7

I. INTRODUCTION

Numerical calculations for quantum many-body systems face fundamental limitations due to the exponential growth of the Hilbert space with an increase in the degrees of freedom. To overcome this problem, diverse numerical approaches have been developed, including quantum Monte Carlo methods [1], density-matrix renormalization groups [2], and tensor networks [3]. In addition to these approaches, novel methods based on artificial neural networks have recently attracted much interest [4–8]. Leveraging the flexibility and adaptability of neural networks for representing various features, many-body wave functions are encoded on neural networks with a comparatively small number of network parameters, which are optimized using machine-learning techniques. The method of neural-network quantum states was first applied to spatially discrete problems, such as spins and particles on lattices [4,9–12], and then extended to continuous-space problems, such as interacting bosons [13–19], fermionic systems [20–27], and nucleons in nuclei [28–34].

Neural-network quantum states are not only powerful tools for ground states, but also useful for exploring excited states [35–40]. An excited state belonging to a symmetry sector different from the ground state can be simply obtained by minimizing the energy by fixing quantum numbers, such as momentum or magnetization. Even if the desired excited state cannot be distinguished from the ground state by quantum numbers, one can still compute the excited state by projecting it to the Hilbert space orthogonal to the ground state. Using these procedures, low-lying excitations of the Heisenberg model, the Bose-Hubbard model [35], and the J_1-J_2 model [36,37] have been calculated. More recently, multiple excited states of benzene-scale molecules have been computed by regarding them as the ground state of an extended system [39].

While neural-network quantum states have been successfully applied to a variety of quantum systems, their application to strongly correlated few-body systems in continuous space has remained less explored, especially for the excited states of three or more particles. Such problems are of pivotal importance in nuclear physics and cold atomic systems. In nuclear physics, a rich variety of nuclear structures and reactions emerge from strong interactions between protons and neutrons [41]. In cold atoms, strongly correlated few- and many-body systems can be realized by tuning the s-wave scattering length between the atoms via a Feshbach resonance [42,43], enabling the experimental observation of the Efimov states [44–51]. The Efimov states are peculiar three-body bound states characterized by Borromean binding and discrete scale invariance, which manifest as a quantum anomaly in the underlying field theory [47–52]. Owing to this discrete scale invariance, successive Efimov states differ greatly in spatial extent, forming a multiscale structure that provides a stringent test for numerical methods. To understand the physical nature of the Efimov states and to gain insight into many-body physics from a few-body perspective, few-body systems with large s-wave scattering lengths have been extensively studied both experimentally [46,53–59] and theoretically [60–65]. These studies have revealed that, in addition to three-body bound states, four-, five-, and six-body states tied to Efimov trimers also appear universally [55–57,61–63,65]. This hierarchy of few-body clusters provides an important step toward a unified understanding of how few-body systems evolve into many-body systems as the particle number increases [62,63,66–69].

As Efimov binding is a highly quantum-mechanical phenomenon that defies classical description, it is an ideal test bed

*Contact author: shimpei.endo@uec.ac.jp
†Contact author: hiroki.saito@uec.ac.jp

for benchmarking novel numerical methods. Indeed, strongly correlated few-body problems have been accurately solved using the correlated Gaussian basis [61,70,71], adiabatic hyperspherical expansion [61], Gaussian expansion [64,65], and hyperspherical harmonics expansion [63]. Recently, neural-network methods have also been applied to continuous-space few-body systems with large s-wave scattering lengths, but so far mainly to ground states: The ground states of few-boson systems have been obtained for short-range model potentials in Ref. [14], and also for $^4\mathrm{He}_N$ atomic clusters in Ref. [18]. Neural-network calculations of excited states, in turn, have been demonstrated mainly for lattice models [35–37] and molecules [38–40], leaving the strongly correlated, spatially extended excited states of Efimov physics essentially unexplored. Since the Efimov states appear not only in cold atoms but also universally in a wide range of systems with large s-wave scattering lengths, such as weakly bound nuclear systems [51,72–74], ^{4}He clusters [75], and magnetic systems [76], establishing a neural-network method capable of describing Efimov physics is useful for exploring few-body phenomena across various fields and scales.

Here, we develop a neural-network method for computing the ground and first excited states of strongly interacting few-body systems in three-dimensional continuous space, and show that it properly captures the Efimov states and their associated few-body clusters. We adopt the neural-network approach of Ref. [14], which computes the ground-state energy with a fully connected feedforward neural network using suitably chosen Jacobi coordinates as inputs. We extend this approach to the first excited states using the orthogonal-projection method [35], and to mass-imbalanced three-body fermionic systems composed of two identical fermions and a third particle. To stabilize the optimization for the spatially extended excited states, we multiply the network output by a two-body correlation factor that accounts for the short-range two-body behavior. Using this method, we obtain the ground and first excited states for three- to six-body systems of identical bosons at unitarity, as well as for the mass-imbalanced fermionic system. The resulting energies are in good agreement with previous results [63,70,71]. For the three-body systems, in particular, our results are consistent with the expected universal features of the Efimov states, such as the discrete scale invariance and the characteristic wave-function structures. For the mass-imbalanced fermionic system, we further find the critical behavior as the mass ratio approaches 13.6 [45,77]. Because our approach relies on simple, physically motivated ingredients (Jacobi coordinates, rotationally invariant inputs, and two-body correlations), the method is straightforward to implement and can be adapted with only minor modifications to a range of strongly correlated few-body problems.

The rest of this paper is organized as follows. Section II explains the method of neural-network quantum states. Sections III and IV study systems of identical bosons and two identical fermions with one particle, respectively. In each of these sections, the method and numerical results are presented. Section V gives the conclusions and suggestions for future research.

II. NEURAL-NETWORK QUANTUM STATES

To represent a wave function, we use a fully-connected feedforward neural network, which consists of an input layer, hidden layers, and an output layer [78]. The input layer receives a real-valued vector,

$$\boldsymbol{u}^{\mathrm{in}} = \left(u_1^{\mathrm{in}}, u_2^{\mathrm{in}}, \ldots, u_{N_{\mathrm{in}}}^{\mathrm{in}}\right), \tag{1}$$

constructed from the particle coordinates, where N_{in} denotes the number of input values. We will specify later how to construct $\boldsymbol{u}^{\mathrm{in}}$ from the particle coordinates in a uniform space. The neural network contains L_h hidden layers. The units in the ℓth hidden layer ($\ell = 1, 2, \ldots, L_h$) are defined recursively as

$$u_i^{(\ell)} = \sum_{j=1}^{N_{\ell-1}} W_{ji}^{(\ell)} f\left(u_j^{(\ell-1)}\right) + b_i^{(\ell)}, \tag{2}$$

where N_ℓ denotes the number of units in the ℓth layer, $W^{(\ell)}$ is a real $N_{\ell-1} \times N_\ell$ matrix, and $b^{(\ell)}$ is a real N_ℓ component vector. The input layer corresponds to $\ell = 0$. As an activation function $f(x)$, we adopt the SiLU (Sigmoid Linear Unit) function:

$$f(x) = \frac{x}{1 + e^{-x}}. \tag{3}$$

We found that this activation function allows more stable and accurate calculations than those obtained using other activation functions, such as the ReLU (Rectified Linear Unit) and sigmoid functions. The output layer gives the following two real values:

$$u_k^{\mathrm{out}} = \sum_{i=1}^{N_{L_h}} W_{ik}^{(L_h+1)} f\left(u_i^{(L_h)}\right) \quad (k = 1, 2). \tag{4}$$

In this paper, we use a neural network with $L_h = 2$ and $N_\ell = 32$. Finally, the network outputs a single value A:

$$A = u_1^{\mathrm{out}} \exp\left(u_2^{\mathrm{out}}\right), \tag{5}$$

so that A can take either sign with exponential nonlinearity. The many-body trial wave function Ψ is constructed using the network output A in combination with functions of X, where $X = \{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\}$ denotes the particle coordinates and W denotes the network parameters, $W^{(\ell)}$ and $b^{(\ell)}$.

According to the variational principle, the expectation value $E(W)$ of the Hamiltonian $\hat{H}$ for the trial wave function Ψ is not less than the true ground-state energy E_0:

$$E(W) = \frac{\int \Psi^* \hat{H} \Psi dX}{\int |\Psi|^2 dX} \geqslant E_0, \tag{6}$$

where $dX = d\boldsymbol{r}_1 \cdots d\boldsymbol{r}_N$. We minimize $E(W)$ with respect to W using the variational Monte Carlo method. In the following sections, we rewrite $E(W)$ in a form generally given by

$$E(W) = \frac{\int F(P, W) dP}{\int |\Psi(P, W)|^2 dP}, \tag{7}$$

where F is a function constructed from the network output A and the Jacobi coordinates $P = \{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots, \boldsymbol{\rho}_{N-1}\}$ of the particle positions, and $dP = d\boldsymbol{\rho}_1 d\boldsymbol{\rho}_2 \cdots d\boldsymbol{\rho}_{N-1}$. By taking the Metropolis-Hastings sampling of the coordinates P with a

probability distribution

$$p(P,W)=\frac{|\Psi(P,W)|^2}{\int|\Psi(P,W)|^2dP}, \tag{8}$$

we can evaluate the multidimensional integration in Eq. (7) as

$$E(W)=\int p(P,W)\frac{F(P,W)}{|\Psi(P,W)|^2}dP\simeq\frac{1}{N_s}\sum_{n=1}^{N_s}\frac{F(P_n,W)}{|\Psi(P_n,W)|^2}, \tag{9}$$

where N_s is the number of samples. Similarly, the gradient of the energy with respect to the network parameters is obtained as

$$\frac{\partial E}{\partial W}\simeq\frac{1}{N_s}\sum_{n=1}^{N_s}\frac{1}{|\Psi|^2}\left(\frac{\partial F}{\partial W}-E\frac{\partial|\Psi|^2}{\partial W}\right). \tag{10}$$

Using this gradient, we update the network parameters using the Adam optimizer [79] to minimize the energy. The learning rate in the Adam optimizer is typically chosen to be 10^{-3} to 10^{-4}. Large learning rates make the optimization process unstable and the network parameters diverge. Once stability is achieved, the final results are insensitive to the learning rate.

We use the projection method to obtain the excited state. Namely, we define the following wave function [35]:

$$\Psi(P,W)=\Psi'(P,W)-\lambda(W)\Psi_0(P,W_0), \tag{11}$$

where $\Psi_0(P,W_0)$ is the ground-state wave function obtained using the above method and $\Psi'(P,W)$ is a wave function constructed from a network that is different from that used for the ground state. The parameter λ in Eq. (11) is given by

$$\lambda(W)=\frac{\int\Psi_0^*(P,W_0)\Psi'(P,W)dP}{\int|\Psi_0(P,W_0)|^2dP}, \tag{12}$$

which is also evaluated using Monte Carlo sampling. The wave function Ψ in Eq. (11) is orthogonal to the ground state Ψ_0 as long as the evaluated value of λ is accurate. By minimizing the energy with respect to the network parameters W, we obtain the first excited state. The gradient of the energy can be calculated using the automatic differentiation provided by a machine-learning framework.

III. BOSONIC SYSTEMS

A. Method

We first apply our neural-network method to systems of N identical bosons with mass m in three-dimensional free space. The position of the ith particle is denoted by the Cartesian coordinate $\boldsymbol{r}_i\in\mathbb{R}^3$. The ith and jth particles interact with each other through the two-body potential $V(r_{ij})$, where $r_{ij}=|\boldsymbol{r}_i-\boldsymbol{r}_j|$. The Hamiltonian for the system is given by

$$\hat{H}=-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\frac{\partial^2}{\partial\boldsymbol{r}_i^2}+\sum_{i<j}V(r_{ij}). \tag{13}$$

The system has translational symmetry and we eliminate the center-of-mass motion using the Jacobi coordinates,

$$\boldsymbol{\rho}_{N-j}=\sqrt{\frac{2j}{j+1}}\left[\boldsymbol{r}_{j+1}-\frac{1}{j}(\boldsymbol{r}_1+\boldsymbol{r}_2+\cdots+\boldsymbol{r}_j)\right],\quad j=1,2,\ldots,N-1, \tag{14}$$

and the center-of-mass coordinate,

$$\boldsymbol{\rho}_{\rm cm}=\sqrt{\frac{2}{N(N+1)}}(\boldsymbol{r}_1+\boldsymbol{r}_2+\cdots+\boldsymbol{r}_N). \tag{15}$$

Transforming from the Cartesian coordinates to $\{\boldsymbol{\rho}_{\rm cm},\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,\ldots,\boldsymbol{\rho}_{N-1}\}$, we can decompose the Hamiltonian $\hat{H}$ as

$$\hat{H}=-\frac{\hbar^2}{m}\frac{1}{N+1}\frac{\partial^2}{\partial\boldsymbol{\rho}_{\rm cm}^2}+\hat{H}', \tag{16}$$

where

$$\hat{H}'=-\frac{\hbar^2}{m}\sum_{j=1}^{N-1}\frac{\partial^2}{\partial\boldsymbol{\rho}_j^2}+\sum_{i<j}V(r_{ij}). \tag{17}$$

The relative distance r_{ij} in the potential terms can only be expressed by $\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,\ldots,\boldsymbol{\rho}_{N-1}$, and therefore, the decomposed part $\hat{H}'$ does not involve $\boldsymbol{\rho}_{\rm cm}$. We therefore only consider $\hat{H}'$ and minimize its expectation value in the following.

The Hamiltonian $\hat{H}'$ has rotational symmetry and here we consider only rotationally symmetric wave functions. Rotationally symmetric functions are functions of the elements of the Gram matrix,

$$G_{ij}=\boldsymbol{\rho}_i\cdot\boldsymbol{\rho}_j\quad(i,j=1,2,\ldots,N-1). \tag{18}$$

Although the number of independent elements of the Gram matrix is $3N-6$, we use all of the $N(N-1)/2$ elements $G_{i\leqslant j}$ as the neural network's inputs in Eq. (1) as

$$\boldsymbol{u}^{\rm in}=(G_{11},G_{12},\ldots,G_{N-1,N-1}). \tag{19}$$

With this choice of inputs, the translational and rotational degrees of freedom are removed at the input level. The neural network can thus efficiently optimize its parameters while keeping the symmetries of the quantum states. On the other hand, the particle-exchange symmetry required for identical bosons is not explicitly imposed. Nevertheless, we numerically find that the obtained wave functions are sufficiently accurate, almost satisfying the particle-exchange symmetry [14], which will be shown in the numerical results in Sec. III B.

To accelerate the convergence, we explicitly incorporate the two-body correlations into the trial wave function. Let $\phi(r)$ be a function expressing the short-range behavior of two particles. We construct the trial wave function Ψ using $\phi(r)$ as

$$\Psi=\varphi A, \tag{20}$$

where

$$\varphi=\prod_{i<j}\phi(r_{ij}) \tag{21}$$

is the product of the two-body correlation $\phi(r)$ for all particle pairs and A is the network output in Eq. (5). The multiplication of the two-body correlation φ assists the trial wave function Ψ in expressing the short-range behavior, by which short-range variations of the network output A can be mitigated. Using the wave function in Eq. (20), the energy integral of the Hamiltonian in Eq. (13) becomes

$$\begin{aligned}
&\int dX\left[-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Psi\frac{\partial^2}{\partial \boldsymbol{r}_i^2}\Psi+\sum_{i<j}V(r_{ij})\Psi^2\right]\\
&=\int dX\left[-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\varphi A\left(2\frac{\partial\varphi}{\partial \boldsymbol{r}_i}\cdot\frac{\partial A}{\partial \boldsymbol{r}_i}+\varphi\frac{\partial^2 A}{\partial \boldsymbol{r}_i^2}\right)+V_{\rm eff}\Psi^2\right]\\
&=\int dX\left[\frac{\hbar^2}{2m}\sum_{i=1}^{N}\varphi^2\left(\frac{\partial A}{\partial \boldsymbol{r}_i}\right)^2+V_{\rm eff}\Psi^2\right]. \qquad (22)
\end{aligned}$$

The effective potential $V_{\rm eff}$ on the second line of Eq. (22) is defined as

$$\begin{aligned}
&-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\frac{\partial^2\varphi}{\partial \boldsymbol{r}_i^2}\\
&=-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\sum_{j\neq i}\left[\frac{1}{\phi(r_{ij})}\frac{\partial^2\phi(r_{ij})}{\partial \boldsymbol{r}_i^2}+\sum_{k\neq i,j}\boldsymbol{b}_{ij}\cdot\boldsymbol{b}_{ik}\right]\varphi\\
&\equiv-\sum_{i<j}V(r_{ij})\varphi+V_{\rm eff}\varphi, \qquad (23)
\end{aligned}$$

where

$$\boldsymbol{b}_{ij}=\frac{1}{\phi(r_{ij})}\frac{\partial\phi(r_{ij})}{\partial \boldsymbol{r}_i}. \qquad (24)$$

Eliminating the center-of-mass coordinate using the Jacobi coordinates, we can rewrite Eq. (22) as

$$\int\Psi\hat{H}'\Psi dP=\int dP\left[\frac{\hbar^2\varphi^2}{m}\sum_{j=1}^{N-1}\left(\frac{\partial A}{\partial \boldsymbol{\rho}_j}\right)^2+V_{\rm eff}\Psi^2\right]. \qquad (25)$$

The expectation value of the energy $\int\Psi\hat{H}'\Psi dP/\int\Psi^2 dP$ is minimized using the method described in Sec. II, where the function F in Eq. (7) corresponds to

$$F(P,W)=\frac{\hbar^2\varphi^2(P)}{m}\sum_{j=1}^{N-1}\left[\frac{\partial A(P,W)}{\partial \boldsymbol{\rho}_j}\right]^2+V_{\rm eff}(P)\Psi^2(P,W). \qquad (26)$$

The sampling of the particle positions is performed using the Metropolis-Hastings algorithm. Initially, the particle positions P in the Jacobi coordinates are set to be random. In each sampling, a random displacement vector δP is generated, which obeys the normal distribution with standard deviation σ. The value of σ/b is typically taken to be 1 for the ground states and 1–20 for the first excited states, where b is a typical length scale of the two-body interaction. If $|\Psi(P+\delta P,W)|^2\geqslant|\Psi(P,W)|^2$, the displacement of the particle position is accepted; otherwise, it is accepted with the probability $|\Psi(P+\delta P,W)|^2/|\Psi(P,W)|^2$. In the early stage of the optimization, samples sometimes diverge to infinity during the random walk. To suppress such unphysical divergence of particle configurations, we add an auxiliary harmonic potential [14],

$$\frac{m\omega^2}{4}\sum_{i=1}^{N-1}\rho_i^2, \qquad (27)$$

to the Hamiltonian with $\omega\sim10^{-1}$ to $10^{-3}\hbar/(mb^2)$. After the first 1000 updates, the auxiliary harmonic potential is removed to obtain the results in free space. The final results are insensitive to the strength and removal schedule of the auxiliary harmonic potential.

In the numerical calculations shown in the next subsection, we employ the Pöschl-Teller potential for the two-body interaction potential,

$$V(r)=-\frac{2\hbar^2}{mb^2}\frac{1}{\cosh^2\left(\frac{r}{b}\right)}, \qquad (28)$$

where b characterizes the range of the interaction. The potential strength $2\hbar^2/mb^2$ is chosen such that the first s-wave two-body bound state is about to appear, corresponding to an infinite s-wave scattering length (i.e., the unitary limit), which provides a suitable condition for exploring Efimov physics. For this potential, the two-body Schrödinger equation can be solved analytically, and the zero-energy solution is obtained as [80,81]

$$\phi(r)=\frac{1}{r}\tanh\frac{r}{b}. \qquad (29)$$

We employ this solution for the two-body correlation ϕ in Eq. (21). The second derivative in Eq. (23) then becomes

$$-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\sum_{j\neq i}\frac{\partial^2\phi(r_{ij})}{\partial \boldsymbol{r}_i^2}=-\sum_{i<j}V(r_{ij})\phi(r_{ij}), \qquad (30)$$

and exactly cancels the potential term of $V(r_{ij})$ on the final line, which yields the simple form of the effective potential as

$$V_{\rm eff}=-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\sum_{j\neq i}\sum_{k\neq i,j}\boldsymbol{b}_{ij}\cdot\boldsymbol{b}_{ik}, \qquad (31)$$

with Eq. (24) being

$$\boldsymbol{b}_{ij}=\left(-\frac{1}{r_{ij}}+\frac{2}{b\sinh\frac{2r_{ij}}{b}}\right)\frac{\boldsymbol{r}_i-\boldsymbol{r}_j}{r_{ij}}. \qquad (32)$$

Thus, the two-body potential is eliminated from the Hamiltonian in Eq. (22) and the effective potential appears instead. For example, in the case of $N=3$, $V_{\rm eff}=-\hbar^2(\boldsymbol{b}_{12}\cdot\boldsymbol{b}_{13}+\boldsymbol{b}_{21}\cdot\boldsymbol{b}_{23}+\boldsymbol{b}_{31}\cdot\boldsymbol{b}_{32})/m$. We note that this cancellation is not specific to the Pöschl-Teller potential, but occurs generally if ϕ is taken to be the exact zero-energy solution of the two-body Schrödinger equation with $V(r_{ij})$. The present formalism is therefore applicable to a broad class of interaction potentials whose two-body Schrödinger equation can be solved accurately, either analytically or numerically. Furthermore, as shown in Sec. IV, the above formalism remains effective even when the exact two-body solution is unavailable: An approximate ϕ can be used to construct $V_{\rm eff}$ according to Eq. (23), despite the absence of the exact cancellation.

In a previous work [14], the ground states for several bosons interacting through the Gaussian potential were calculated using neural-network quantum states without the two-body correlations φ. As we will show in the next subsection, the inclusion of φ is crucial for increasing the accuracy and stability of the calculations for practical use, which enables the study of the excited states.

B. Results

We present numerical results for the ground and first excited states of N-boson systems with $N = 3$, 4, 5, and 6 interacting via the unitary Pöschl-Teller potential. To assess the accuracy of the present method, we compare our results with those obtained using the hyperspherical harmonic expansion method [63], where the same Hamiltonian with the unitary Pöschl-Teller potential is used. In all the calculations presented in Secs. III and IV, 10^5 Monte Carlo samples are used in each update step of the neural network to evaluate the gradient in Eq. (10). To accelerate the sampling process using graphics processing units, these 10^5 samples are taken in parallel with an interval of 16 Metropolis steps. This interval is sufficient, since the final results remain unchanged when it is increased to, e.g., 32.

Figure 1 shows the convergence behavior of the ground-state energy with respect to the update steps. In Fig. 1(a), the network output in Eq. (5) is directly used as the wave function and the Hamiltonian in Eq. (17) is evaluated. In Fig. 1(b), the two-body correlation φ is used as in Eq. (20) and the energy in Eq. (25) is evaluated. For both methods, the N-body ground-state energies converge to the values in Ref. [63], demonstrating that the neural networks can accurately describe the N-body Borromean clusters associated with the Efimov states. A comparison between Figs. 1(a) and 1(b) clearly shows that the inclusion of the two-body correlation φ significantly improves the stability and smoothness of the convergence, while noticeable fluctuations are observed without φ. This improvement originates from the more accurate description of the short-range two-body correlations by φ, by which the network does not need to represent the sharp short-range profile of the wave function. As explained later, the posttrained ground-state energy for three bosons is evaluated to be $-0.134\,71(2)$ with the two-body correlation φ. In contrast, the posttrained energy is evaluated to be $-0.1346(2)$ without φ, with an uncertainty 10 times larger.

As mentioned in Sec. III A, the particle-exchange symmetry required for identical bosons is not explicitly imposed in this method. To check whether the obtained states satisfy the particle-exchange symmetry, we evaluate the overlap integral defined by [14]

$$S = \frac{2}{N(N-1)} \sum_{j<k} \frac{\left(\int \Psi_{jk} \Psi dX\right)^2}{\int \Psi_{jk}^2 dX \int \Psi^2 dX}, \tag{33}$$

where Ψ_{jk} is the wave function for which the positions $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$ of the jth and kth particles are exchanged. When the wave function Ψ exactly satisfies the particle-exchange symmetry, S is equal to unity. The dashed lines in Fig. 1 show the values of S evaluated by the Monte Carlo sampling, which are larger than 0.999 for both panels (a) and (b), indicating

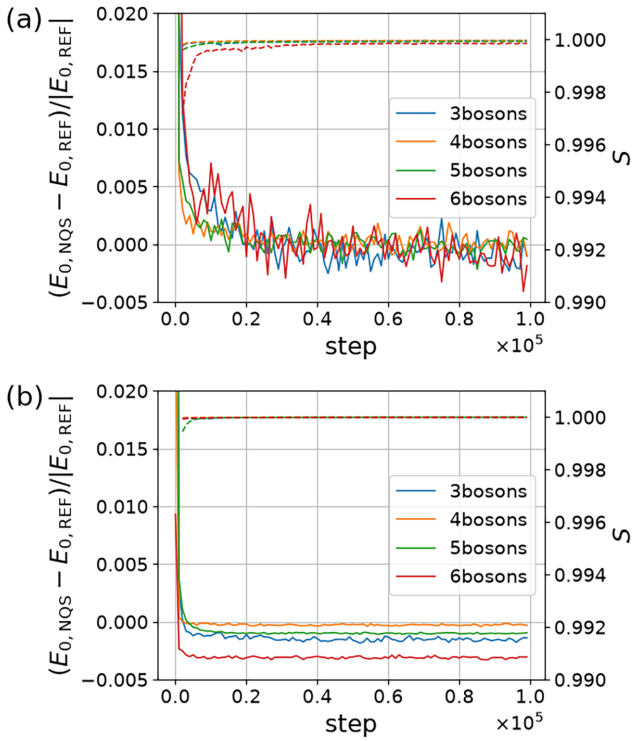

FIG. 1. Optimization process of neural networks for N-boson ground states. The solid lines show the relative difference between the obtained ground-state energy $E_{0,\mathrm{NQS}}$ and the reference energy $E_{0,\mathrm{REF}}$ taken from Ref. [63], and the dashed lines show the overlap integral S defined in Eq. (33). The horizontal axis represents the number of optimization steps. (a) Two-body solution is not included in the trial wave function: Equation (5) is used as the wave function and the expectation value of the Hamiltonian in Eq. (17) is evaluated. (b) Two-body solution is included in the trial wave function: The expression in Eq. (25) with Eq. (31) is evaluated with the wave function in Eqs. (20) and (21) with Eq. (29). In both cases, 10^5 Monte Carlo samples are used in each step, and the values of the energies and S are averaged over every 1000 update steps.

that the obtained wave functions satisfy the particle-exchange symmetry to a good approximation.

Using the obtained ground-state energies and wave functions, we next compute the first excited states. To avoid the numerical instability originating from the small binding energies and large spatial extent of excited few-body clusters [61–63,65], an auxiliary harmonic potential in Eq. (27) with $\omega = 10^{-3}\hbar/(mb^2)$ is applied during the first 1000 steps. In the following, we only present the results obtained with the two-body correlation φ, since the optimization process is unstable without φ. Figure 2 shows the convergence behavior of the energies of the first excited states. For all N, the excited-state energies converge to negative values, indicating the formation of bound states. Although the fluctuations in the energies in Fig. 2(b) for $N = 4$–6 are $\sim$1%, those in Fig. 2(a) for $N = 3$ are $\sim$10%. The large fluctuation for $N = 3$ is due to the extended spatial structure of the Efimov state; the spatial scale of the excited state is $\simeq$20 times larger than that of the ground state, while the short-range node structure is also important for ensuring orthogonality with the ground state. This multi-scale property of the Efimov states makes the Monte Carlo sampling less efficient. For $N = 4$–6, in contrast, the first

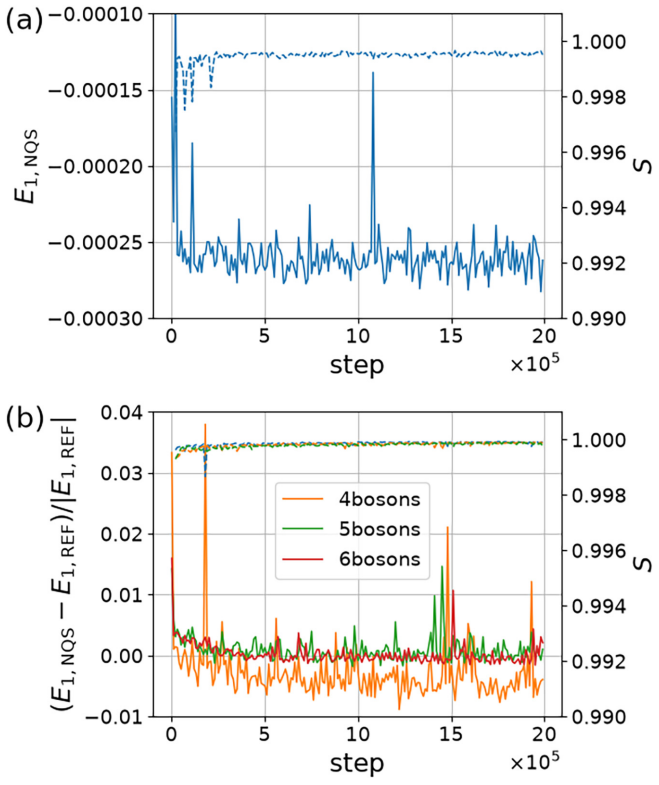

FIG. 2. Optimization process of neural networks for first excited states of N-boson systems. The solid and dashed lines show the energies and overlap integral S, respectively. (a) $N = 3$. Energy is normalized by $\hbar^2/(mb^2)$. (b) $N = 4$, 5, and 6, where the values of $E_{1,\mathrm{REF}}$ are taken from Ref. [63]. Since the excited-state energy for $N = 3$ is not explicitly provided in Ref. [63], only $E_{1,\mathrm{NQS}}$ is plotted in panel (a). In panels (a) and (b), 10^5 Monte Carlo samples are used in each step, and the values of the energies and S are averaged over every 10^4 update steps.

excited N-body state is tied below the $(N-1)$-body ground state [48–50,61], and therefore remains relatively compact. Consequently, the fluctuations arising from the Monte Carlo sampling are less significant. The dashed lines in Fig. 2 show the values of S defined in Eq. (33), which are larger than 0.999 for all the cases, indicating that the bosonic particle-exchange symmetry is well satisfied also for the first excited states.

Using the trained neural networks, we next determine the energies of the ground and first excited states. In this posttrained evaluation, each energy is obtained from 10^8 (10^9) Monte Carlo samples, followed by 10^3 (10^4) additional parameter updates of the neural network, for the ground states (first excited states). This procedure is repeated 10 times and we obtain 10 energy estimates. In addition, we repeat the entire procedure (training and posttrained evaluation) 5 times using different random seeds. Thus, we obtain 50 energy estimates in total, and we take their average and standard deviation. The results are summarized in Table I, where they are compared with the values in Ref. [63]. Our values are in good agreement with those in Ref. [63]. For $N = 3$, the ratio $\sqrt{E_0/E_1}$ is consistent with the universal value predicted from the zero-range theory, $\simeq$22.7. This agreement is expected because the finite-range effect is of order $(r_{\mathrm{eff}}\kappa_0)^2$ [82], where $\kappa_0 = \sqrt{m|E_0|}/\hbar$ and the effective range is $r_{\mathrm{eff}} = 2b$

TABLE I. Ground-state energies E_0, first-excited-state energies E_1, and ratios $\sqrt{E_0/E_1}$ for N-boson systems. The value of 22.4 for $N = 3$ is taken from the main text of Ref. [63] and the other reference values are taken from Table II in Ref. [63]. The energies are normalized by $\hbar^2/(mb^2)$.

	E_0	E_1	$\sqrt{E_0/E_1}$
$N = 3$ (REF)	-0.1345		22.4
$N = 3$ (NQS)	$-0.13471(2)$	$-2.63(9)\times10^{-4}$	22.6(4)
$N = 4$ (REF)	-0.8259	-0.1548	2.31
$N = 4$ (NQS)	$-0.82612(6)$	$-0.1554(4)$	2.305(3)
$N = 5$ (REF)	-2.313	-0.9428	1.57
$N = 5$ (NQS)	$-2.31531(9)$	$-0.9421(18)$	1.568(2)
$N = 6$ (REF)	-4.731	-2.667	1.33
$N = 6$ (NQS)	$-4.7456(4)$	$-2.663(8)$	1.335(2)

for the unitary Pöschl-Teller potential, and thus $r_{\mathrm{eff}}\kappa_0 = 0.73$. Because $E_0/E_1 \simeq 500$ for $N = 3$, even a small admixture of the ground state into the first excited state can produce a noticeable error in E_1. The relative uncertainty of order 10^{-4} in E_0 suggests a residual ground-state admixture into the first excited state of the same order, through Eq. (11), leading to fluctuations in E_1 of order $\sim 10^{-4}E_0$, which is consistent with the uncertainty in E_1 of order 10^{-5} reported in Table I.

We have checked the statistical properties of the Monte Carlo samples. The acceptance rate of the proposals in the Metropolis-Hastings algorithm is typically 30%–50%, which depends on the displacement σ in the random walk of the Jacobi coordinates; a larger σ results in a smaller acceptance rate. The autocorrelation function of the sampled Jacobi coordinates decays exponentially with respect to the steps of the random walk. The typical $1/e$ decay time is 10–30 steps. This number of decay steps justifies the approximate independence between subsequent samples taken with an interval of 16 steps.

Figures 3(a) and 3(b) show the probability distributions $P(R)$ for the ground and first excited states, respectively, of the Efimov system ($N = 3$). Here, $P(R)$ is obtained as a histogram of the hyperradius $R = \sqrt{\rho_1^2 + \rho_2^2}$, which characterizes the spatial extent of the three-body system. In both states, the distribution $P(R)$ exhibits a pronounced peak at finite R and decays at large R, indicating the formation of bound states. The two distributions have similar shapes and the ratio between the peak positions is $\simeq$15. This larger deviation from the zero-range prediction $\simeq$22.7 reflects the finite-range effect of order $r_{\mathrm{eff}}\kappa_0$. In the first excited state in Fig. 3(b), a node appears in the short-range region, making the state orthogonal to the ground state.

The top-right insets in Fig. 3 show the spatial distribution of the three-body wave function, namely, the probability distribution for the position of the third particle when the other two particles are fixed at the positions denoted by the crosses. The results show the same qualitative features as those observed for ^{4}He trimers [75]: In contrast to the probability density for the ground state [Fig. 3(a)], which forms a compact bondinglike orbital, that for the first excited state [Fig. 3(b)] is concentrated around each particle, reflecting the origin of the Efimov trimer as arising from the exchange of a particle between the other two particles in a classically forbidden region.

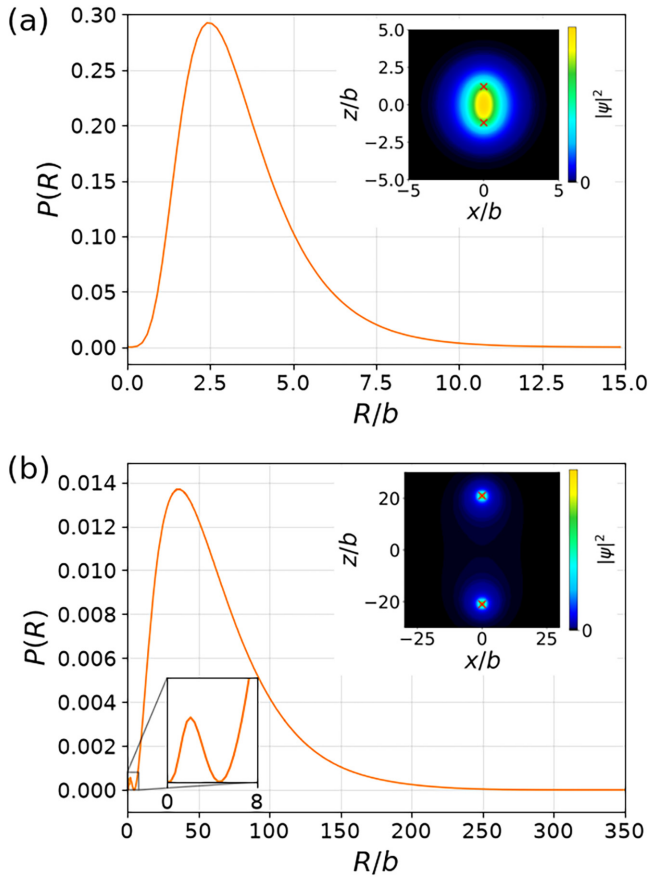

FIG. 3. Probability distribution $P(R)$ as a function of hyperradius $R = \sqrt{\rho_1^2 + \rho_2^2}$ for bosonic three-body system: (a) ground state and (b) first excited state. $P(R)$ is generated from the wave function Ψ as a histogram of samples R. The top-right insets show the probability distribution for the position of the third particle when the other two particles are fixed at the positions indicated by the crosses. The bottom-left inset in panel (b) magnifies the line near the origin.

The above results are obtained using the fully connected feedforward neural network with two hidden layers containing 32 hidden units in each layer. We also examined a network with 64 hidden units in each layer and a network with three hidden layers containing 32 hidden units each. We calculated the ground-state energy of three identical bosons using these networks and obtained $E_0 = -0.134\,71(2)$ for both cases, which is the same as that reported in Table I. Exploring alternative neural-network architectures is an interesting direction for future work.

IV. SYSTEM OF TWO IDENTICAL FERMIONS AND ONE PARTICLE

In this section, we apply the neural-network method to a three-body system composed of two identical fermions (mass M) and a third particle (mass m) interacting via unitary interspecies interactions. While the equal-mass system does not exhibit any three-body bound states due to the Pauli repulsion between the identical fermions, it is dominated by the attraction mediated by the third particle if the mass ratio exceeds the critical value, $M/m \gtrsim 13.6$ [45,70,71,77]. Such highly mass-imbalanced fermionic systems have recently been realized in cold atom mixtures of Er-Li [83] and Dy-Li [84], where the fermionic Efimov states are expected to appear in the $\ell^{\Pi} = 1^-$ angular-momentum and parity sector [85,86].

A. Method

We consider a three-body system composed of two identical fermions located at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ with mass M, and another particle located at $\boldsymbol{r}_3$ with mass m. The two-body interaction is introduced only between each fermion and the third particle; that is, no interaction is assumed between the two identical fermions, since there is no s-wave interaction between them. The Hamiltonian for this system is given by

$$\hat{H}_F = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial \boldsymbol{r}_1^2} + \frac{\partial^2}{\partial \boldsymbol{r}_2^2}\right) - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial \boldsymbol{r}_3^2} + V(r_{13}) + V(r_{23}). \tag{34}$$

Using the Jacobi coordinates,

$$\boldsymbol{\rho}_1 = \sqrt{\frac{M+m}{2m}}(\boldsymbol{r}_2 - \boldsymbol{r}_1), \tag{35a}$$

$$\boldsymbol{\rho}_2 = \sqrt{\frac{2(M+m)}{2M+m}}\left(\boldsymbol{r}_3 - \frac{\boldsymbol{r}_1 + \boldsymbol{r}_2}{2}\right), \tag{35b}$$

we can eliminate the center-of-mass motion from the Hamiltonian and obtain

$$\hat{H}_F' = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial \boldsymbol{\rho}_1^2} - \frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial \boldsymbol{\rho}_2^2} + V(r_{13}) + V(r_{23}), \tag{36}$$

where $\mu = Mm/(M+m)$ is the reduced mass, $r_{13} = |\sqrt{m}\boldsymbol{\rho}_1 + \sqrt{2M+m}\boldsymbol{\rho}_2|/\sqrt{2(M+m)}$, and $r_{23} = |\sqrt{m}\boldsymbol{\rho}_1 - \sqrt{2M+m}\boldsymbol{\rho}_2|/\sqrt{2(M+m)}$.

The Hamiltonian in Eq. (36) commutes with the angular-momentum operator:

$$\hat{L} = \sum_{j=1}^{2}\left(\boldsymbol{\rho}_j \times \frac{\hbar}{i}\frac{\partial}{\partial \boldsymbol{\rho}_j}\right). \tag{37}$$

The eigenstates of the Hamiltonian $\hat{H}_F'$ can therefore be characterized by angular-momentum quantum numbers. We express the Jacobi vectors $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ by their lengths ρ_1 and ρ_2, the angle θ between them, and their overall rotation by the Euler angles ϕ, θ', and ϕ' as

$$\boldsymbol{\rho}_1 = R_z(\phi')R_y(\theta')R_z(\phi)\begin{pmatrix}0\\0\\\rho_1\end{pmatrix} = \rho_1\begin{pmatrix}\sin\theta'\cos\phi'\\\sin\theta'\sin\phi'\\\cos\theta'\end{pmatrix}, \tag{38}$$

$$\boldsymbol{\rho}_2 = R_z(\phi')R_y(\theta')R_z(\phi)\begin{pmatrix}\rho_2\sin\theta\\0\\\rho_2\cos\theta\end{pmatrix} = \rho_2\begin{pmatrix}\cos\theta\sin\theta'\cos\phi' + \sin\theta(\cos\phi\cos\theta'\cos\phi' - \sin\phi\sin\phi')\\\cos\theta\sin\theta'\sin\phi' + \sin\theta(\cos\phi\cos\theta'\sin\phi' + \sin\phi\cos\phi')\\\cos\theta\cos\theta' - \sin\theta\cos\phi\sin\theta'\end{pmatrix}, \tag{39}$$

where $R_{x,y,z}(\alpha)$ is the rotation matrix, with rotation by an angle α around the x, y, or z axis. In this coordinate system, the operator $\hat{\boldsymbol{L}}^2$ can be represented only by the Euler angles (it does not include ρ_1, ρ_2, and θ) [87]. We denote the eigenvalues of $\hat{\boldsymbol{L}}^2$ and $\hat{L}_z = \hbar\partial/(i\partial\phi')$ as $\hbar^2\ell(\ell+1)$ and $\hbar m_z$, respectively, and the corresponding eigenfunction as $Y_{n,\ell,m_z}(\phi,\theta',\phi')$, where the index n identifies the eigenstates within each angular-momentum subspace. The energy is degenerate with respect to m_z due to the rotational symmetry and we only consider the $m_z = 0$ states. We numerically confirmed that the $\ell = 1$ subspace with odd parity has the lowest energy, consistent with the zero-range results, and therefore focus on this channel throughout this section. We note, however, that our formalism can be systematically extended to other angular-momentum and parity channels. The parity transformation, $\boldsymbol{\rho}_1 \to -\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2 \to -\boldsymbol{\rho}_2$, is rewritten as $\theta' \to \pi - \theta'$, $\phi \to \pi - \phi$, and $\phi' \to \phi' + \pi$. There are nine eigenfunctions with $\ell = 1$ [87]. Among them, the eigenfunctions with $m_z = 0$ and odd parity are given by

$$Y_{1,1,0} = \cos\theta', \quad Y_{2,1,0} = \sin\theta'\cos\phi. \tag{40}$$

The general form of the wave function can therefore be written as

$$\psi(\rho_1,\rho_2,\theta;\theta',\phi) = f_1(\rho_1,\rho_2,\theta)\cos\theta' + f_2(\rho_1,\rho_2,\theta)\sin\theta'\cos\phi, \tag{41}$$

where f_1 and f_2 are unknown functions.

We express the functions f_1 and f_2 by two different neural networks as $f_1 = A(P, W_1)$ and $f_2 = A(P, W_2)$, respectively, where A is the network output given in Eq. (5). The wave function must be antisymmetric with respect to the exchange of the two identical fermions. In the Jacobi coordinates, the exchange of the fermions is expressed as $\boldsymbol{\rho}_1 \to -\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2 \to \boldsymbol{\rho}_2$, which are rewritten as $\theta \to \pi - \theta$, $\theta' \to \pi - \theta'$, $\phi \to -\phi$, and $\phi' \to \phi' + \pi$, as found from Eqs. (38) and (39). Thus, in a manner similar to Eqs. (20) and (21), we construct the trial wave function as

$$\Psi = \varphi A_{\text{asym}}, \tag{42}$$

where

$$\varphi = \phi(r_{13})\phi(r_{23}) \tag{43}$$

is the product of two-body correlations and

$$A_{\text{asym}} = \psi(\rho_1,\rho_2,\theta,\theta',\phi) - \psi(\rho_1,\rho_2,\pi-\theta,\pi-\theta',-\phi) \tag{44}$$

ensures antisymmetrization for the fermions. As in the bosonic case in Eq. (23), we can define the effective potential V_{eff} arising from the derivative of φ. The energy integral is written as

$$\int \Psi\hat{H}'_F\Psi dP = \int dP\left[\frac{\hbar^2}{2\mu}\sum_{j=1}^{2}\varphi^2\left(\frac{\partial A_{\text{asym}}}{\partial\boldsymbol{\rho}_j}\right)^2 + V_{\text{eff}}\Psi^2\right], \tag{45}$$

where $dP = d\boldsymbol{\rho}_1 d\boldsymbol{\rho}_2$. The procedures used to optimize the network parameters are the same as those for the bosonic case. The first excited state is searched within the same subspace as the ground state, namely, the $\ell = 1$ and odd parity subspace.

In the following numerical calculations, we use the Pöschl-Teller potential for the two-body interaction potential as

$$V_{\text{PT}}(r) = -\frac{\hbar^2}{\mu b^2}\frac{1}{\cosh^2\frac{r}{b}}, \tag{46}$$

whose s-wave scattering length is infinite, corresponding to the unitary limit. We use the zero-energy analytical solution of this potential in Eq. (29) as the two-body correlation ϕ in Eq. (43). The two-body potential term in the Hamiltonian vanishes, as in the bosonic case, and the effective potential simplifies to

$$V_{\text{eff}} = -\frac{\hbar^2}{m}\boldsymbol{b}_{13}\cdot\boldsymbol{b}_{23}, \tag{47}$$

where $\boldsymbol{b}_{ij}$ is defined in Eq. (32).

B. Results

First, we focus on the mass ratio $M/m = 27.752$ corresponding to the ^{167}Er-^{6}Li system. A mixture of these ultracold atomic gases has been realized experimentally, in which interspecies Feshbach resonances were observed [83]. By fine-tuning the s-wave scattering length to be large and achieving the unitary limit, the Efimov states involving identical fermions are expected to appear in the $\ell = 1$ channel when the mass ratio is larger than the critical value, $M/m \gtrsim 13.6$. While a strong dipole-dipole interaction between the Er atoms may affect the physical properties of the Efimov states of Er-Er-Li, we neglect this interaction in this work and only consider the short-range interactions between Er and Li, tuned to the unitary limit, and demonstrate that our neural-network method is useful for investigating fermionic Efimov physics.

Figure 4 shows the convergence of the energies with respect to the update steps. For both the ground and first excited states, the energies converge to negative values, indicating the formation of bound states. The excited state exhibits larger statistical fluctuations than those for the ground state, reflecting its extended spatial structure, which makes the Monte Carlo integration more difficult, as in the bosonic case. After the convergence of the optimization, we perform the post-trained evaluation of the energies in a manner similar to the bosonic case, giving $E_0 = -1.88(2)\times10^{-2}\hbar^2/(mb^2)$ for the ground state and $E_1 = -4.4(1)\times10^{-4}\hbar^2/(mb^2)$ for the first excited state. The ratio $\sqrt{|E_0/E_1|} \simeq 6.5(1)$ is in reasonable agreement with the universal discrete scale factor predicted from the zero-range theory, $\sqrt{|E_0/E_1|} \simeq 7.19$ [45,47–50,77] for $M/m = 27.752$. The deviation of $\sqrt{|E_0/E_1|} \simeq 6.5$ from 7.19 is attributed to the finite-range effects of the potential and the limited validity of the low-energy condition, which often deteriorate the universal description of the ground Efimov trimer.

The acceptance rates in the Metropolis-Hastings algorithm are similar to those in the bosonic case, i.e., 30%–50%. The autocorrelation function of the samples of $\boldsymbol{\rho}_2$ in Eq. (35b), i.e., the Jacobi coordinates related to the third particle, is also similar to that in the bosonic case; i.e., it decays in $\sim$10 steps. However, the autocorrelation time for $\boldsymbol{\rho}_1$ in Eq. (35a), i.e., the relative coordinate between the two fermions, is much longer: $\sim10^3$ steps. This is because the wave function has a node between the two fermions and the random walk rarely

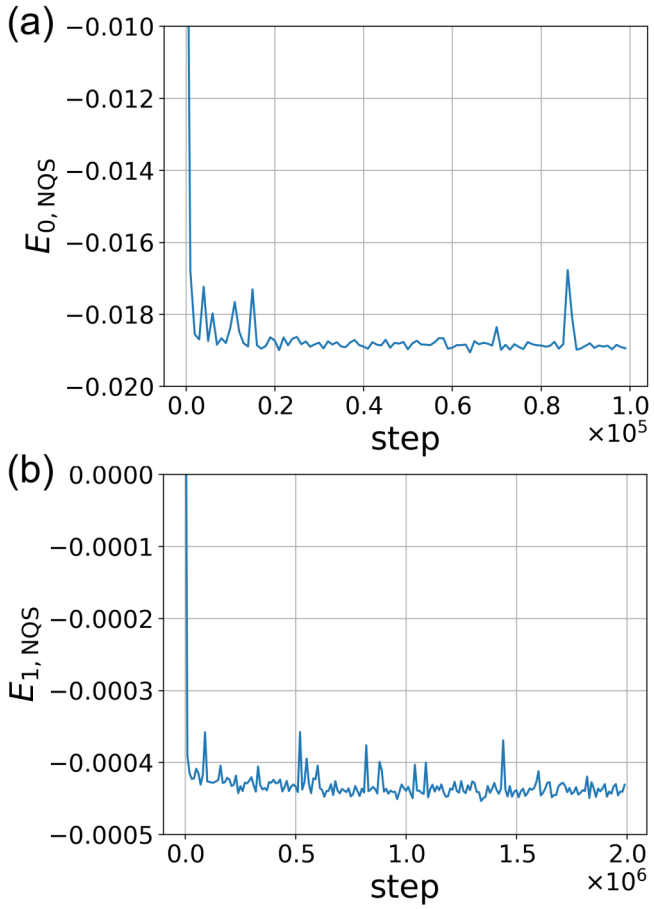

FIG. 4. Optimization process of neural networks for system of two identical fermions and another particle with mass ratio $M/m = 27.752$. (a) Ground-state energy and (b) first-excited-state energy, normalized by $\hbar^2/(mb^2)$. 10^5 Monte Carlo samples are taken in each update. Auxiliary harmonic potential in Eq. (27) is applied during the first 1000 steps, with $\omega = 10^{-2}\hbar/(mb^2)$ for the ground state and $\omega = 10^{-3}\hbar/(mb^2)$ for the excited state.

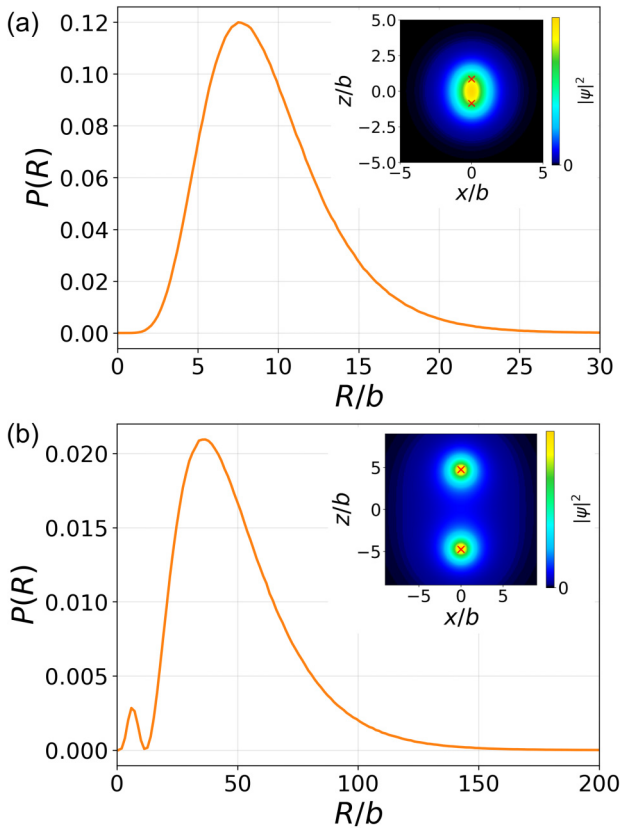

FIG. 5. Probability distribution $P(R)$ as a function of hyperradius R for system of two identical fermions and one particle with mass ratio $M/m = 27.752$: (a) ground state and (b) first excited state. The insets show the probability distributions for the position of the third particle when the two identical fermions are fixed at the positions indicated by the crosses.

exchanges their positions across the node. This is not problematic because the fermion exchange is included in Eq. (44) and it does not need to occur in the sampling. In fact, the autocorrelation times of quantities that are symmetric under fermion exchange, such as ρ_2 and $\sin\theta'$, are $\sim$10 steps.

Figures 5(a) and 5(b) show the probability distributions $P(R)$ with respect to the hyperradius R for the ground and first excited states, respectively. As in the bosonic case, the two distributions have similar shapes with different spatial scales, exhibiting the discrete scale invariance in Efimov physics. The ratio between the peak positions is $\simeq$4.7, which is smaller than the zero-range prediction of $\simeq$7.19 due to the finite-range effects, as in the bosonic case in Fig. 3. In the first excited state [Fig. 5(b)], a node appears in the short-range region, making the state orthogonal to the ground state. The insets in Fig. 5 show the probability distribution for the position of the third particle when the two identical fermions are fixed. The results show the same qualitative features as those for the bosonic three-body systems in Fig. 3.

We also perform the calculation for variable mass ratios and obtain the dependence of the ground-state energy on the mass ratio M/m, as shown by the circles in Fig. 6. As expected from the zero-range Efimov theory [45,47–50], the binding energy increases monotonically with M/m. Extrapolating the data in Fig. 6, we find that the bound state disappears at $M/m \simeq$ 13–14, which is consistent with the zero-range prediction of $\simeq$13.6 [45,77].

We finally consider the case in which the two-body interaction potential is not a Pöschl-Teller potential. We consider the Gaussian potential

$$V_G(r_{ij}) = V_0 e^{-r_{ij}^2/b^2}, \tag{48}$$

where b characterizes the interaction range. The value of V_0 is taken to be $-1.342\hbar^2/(\mu b^2)$, corresponding to the depth at which the two-body bound state is about to appear and hence to the unitary limit. Although the analytical two-body solution is unavailable for the Gaussian potential, we can still perform the neural-network calculation by using the two-body correlation of the Pöschl-Teller potential, exploiting the similar, though not identical, short-range correlations of the two potentials. More specifically, in solving the problem with the Gaussian potential, we use the trial wave function in Eqs. (42) and (43) with the Pöschl-Teller two-body solution in Eq. (29).

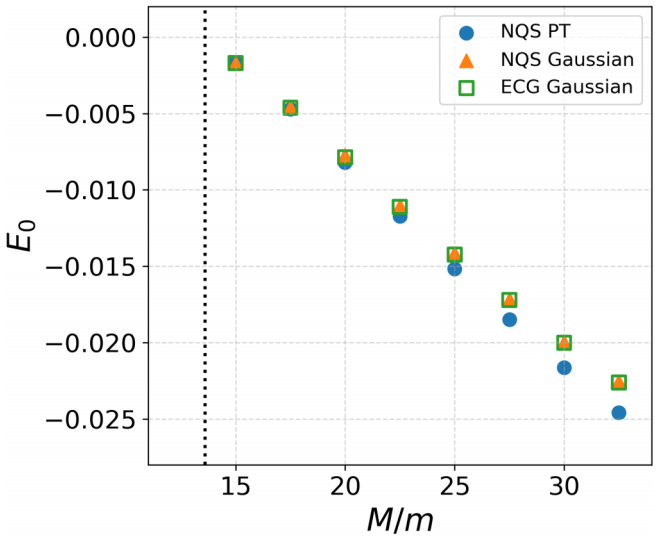

FIG. 6. Ground-state energy [normalized by $\hbar^2/(mb^2)$] of system of two identical fermions and one particle as a function of mass ratio M/m. Circles show the results obtained using the neural-network method for the Pöschl-Teller potential at the unitary limit in Eq. (46). Triangles show the results obtained using the neural-network method for the Gaussian potential in Eq. (48) at the unitary limit. Squares show the results obtained using the explicitly-correlated Gaussian-basis method for the same Gaussian potential [70,71]. The vertical dotted line indicates the critical mass ratio, $M/m \simeq 13.6$. The statistical uncertainties are smaller than the size of the plotted symbols.

The effective potential changes from Eq. (47) to

$$V_{\rm eff} = -\frac{\hbar^2}{m}\boldsymbol{b}_{13}\cdot\boldsymbol{b}_{23} + V_G(r_{13}) + V_G(r_{23}) - V_{\rm PT}(r_{13}) - V_{\rm PT}(r_{23}). \quad (49)$$

The obtained ground-state energies are shown by the triangles in Fig. 6. They are in excellent agreement with those obtained using the explicitly-correlated Gaussian-basis method [70,71] for the same two-body Gaussian potential (squares in Fig. 6). This agreement is attributed to the ability of the neural network to compensate for the relatively small differences between the short-range two-body correlations of the Pöschl-Teller and Gaussian potentials, thereby yielding accurate energies. The above results suggest that the neural-network method is not limited to analytically solvable two-body potentials, but is applicable to problems with a wide variety of interaction potentials.

V. CONCLUSIONS AND DISCUSSION

We developed the method of neural-network quantum states for exploring Efimov physics in few-body quantum systems in three-dimensional continuous space. By using the two-body solution to efficiently incorporate the strong two-body correlations at the unitary limit into the variational wave function, the convergence and accuracy are significantly improved. For systems of N identical bosons, as well as the mass-imbalanced three-body fermionic system, the ground and first excited states are obtained, whose energies agree well with those obtained by previous methods. The ratio between the ground- and first-excited-state energies is found to be consistent with the values predicted by the universal Efimov theory. In addition, our method reproduces the characteristic shapes of the Efimov states' wave functions and captures the disappearance of the Efimov state as the mass ratio approaches the critical value.

In most of the numerical calculations presented in this paper, we used the Pöschl-Teller potential, for which the analytical form of the two-body correlation ϕ and its derivative $\boldsymbol{b}_{ij}$ are available, which accelerates the convergence. However, the proposed method is not restricted to analytically solvable potentials. Indeed, leveraging the fact that the two-body correlations ϕ exhibit similar, though not identical, short-range behavior, we used the two-body correlation of the Pöschl-Teller potential to study a three-body system with a Gaussian potential. The neural networks are found to compensate for the small differences in the two-body correlations of different potentials and to converge rather efficiently to the values obtained using an established method [70,71]. Alternatively, one may employ the numerically obtained ϕ and $\boldsymbol{b}_{ij}$ with appropriate interpolation. These schemes may allow us to solve few-body problems with general interaction potentials, including regimes away from the unitary limit considered here, or dipolar types of interactions whose long-range and anisotropic nature leads to exotic few-body clusters [85,86,88–91].

A major advantage of the present approach, compared with other established methods for few-body studies [61,63–65,70,71], is its simplicity of implementation combined with high flexibility. As seen in our codes (FewBodyNet.py) provided in the Supplemental Material [92], the basic structure of the codes is common to all the systems considered. They are written using the PyTorch framework with the standard machine-learning routines. The codes can be readily extended to a wide variety of systems, such as those with different interaction potentials, harmonically trapped systems, or reduced dimensionality, with only minor modifications. Three- and higher-body interactions are not included in the present study, but can be incorporated straightforwardly. Our method may therefore be useful not only as a numerical tool for benchmarking other theoretical approaches, but also for exploratory studies of previously unexplored few-body systems.

ACKNOWLEDGMENTS

We thank D. Blume for providing the data shown as squares in Fig. 6, which were calculated using the same methods as used in Refs. [70,71]. We thank R. Furuta for valuable suggestions that improved our numerical codes. H.S. was supported by JSPS KAKENHI Grants No. JP23K03276 and No. JP26K00638. S.E. acknowledges support from Matsuo Foundation and JSPS KAKENHI Grant No. JP25K00217.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev. Mod. Phys. **73**, 33 (2001).

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

[3] S.-J. Ran, E. Tirrito, C. Peng, X. Chen, L. Tagliacozzo, G. Su, and M. Lewenstein, *Tensor Network Contractions: Methods and Applications to Quantum Many-Body Systems*, Lecture Notes in Physics, Vol. 964 (Springer International Publishing, Cham, 2020).

[4] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[5] Z.-A. Jia, B. Yi, R. Zhai, Y.-C. Wu, G.-C. Guo, and G.-P. Guo, Quantum neural network states: A brief review of methods and applications, Adv. Quantum Technol. **2**, 1800077 (2019).

[6] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, Restricted Boltzmann machines in quantum physics, Nat. Phys. **15**, 887 (2019).

[7] M. Medvidović and J. R. Moreno, Neural-network quantum states for many-body physics, Eur. Phys. J. Plus **139**, 631 (2024).

[8] H. Lange, A. Van de Walle, A. Abedinnia, and A. Bohrdt, From architectures to applications: A review of neural quantum states, Quantum Sci. Technol. **9**, 040501 (2024).

[9] H. Saito, Solving the Bose-Hubbard model with machine learning, J. Phys. Soc. Jpn. **86**, 093001 (2017).

[10] 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).

[11] Z. Cai and J. Liu, Approximating quantum many-body wave functions using artificial neural networks, Phys. Rev. B **97**, 035116 (2018).

[12] H. Saito and M. Kato, Machine learning technique to find quantum many-body ground states of bosons on a lattice, J. Phys. Soc. Jpn. **87**, 014001 (2018).

[13] M. Ruggeri, S. Moroni, and M. Holzmann, Nonlinear network description for many-body quantum systems in continuous space, Phys. Rev. Lett. **120**, 205302 (2018).

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

[15] G. Pescia, J. Han, A. Lovato, J. Lu, and G. Carleo, Neural-network quantum states for periodic systems in continuous space, Phys. Rev. Res. **4**, 023138 (2022).

[16] J. M. Martyn, K. Najafi, and D. Luo, Variational neural-network ansatz for continuum quantum field theory, Phys. Rev. Lett. **131**, 081601 (2023).

[17] T. Naito, H. Naito, and K. Hashimoto, Multi-body wave function of ground and low-lying excited states using unornamented deep neural networks, Phys. Rev. Res. **5**, 033189 (2023).

[18] W. Freitas and S. A. Vitiello, Synergy between deep neural networks and the variational Monte Carlo method for small $^4\mathrm{He}_N$ clusters, Quantum **7**, 1209 (2023).

[19] P. F. Bedaque, H. Kumar, and A. Sheng, Neural network solutions of bosonic quantum systems in one dimension, Phys. Rev. C **109**, 034004 (2024).

[20] J. Han, L. Zhang, and W. E., Solving many-electron Schrödinger equation using deep neural networks, J. Comput. Phys. **399**, 108929 (2019).

[21] K. Choo, A. Mezzacapo, and G. Carleo, Fermionic neural-network states for *ab-initio* electronic structure, Nat. Commun. **11**, 2368 (2020).

[22] 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).

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

[24] J. W. T. Keeble, M. Drissi, A. Rojo-Francás, B. Juliá-Díaz, and A. Rios, Machine learning one-dimensional spinless trapped fermionic systems with neural-network quantum states, Phys. Rev. A **108**, 063320 (2023).

[25] M. Wilson, S. Moroni, M. Holzmann, N. Gao, F. Wudarski, T. Vegge, and A. Bhowmik, Neural network ansatz for periodic wave functions and the homogeneous electron gas, Phys. Rev. B **107**, 235139 (2023).

[26] E. M. Nordhagen, J. M. Kim, B. Fore, A. Lovato, and M. Hjorth-Jensen, Efficient solutions of fermionic systems using artificial neural networks, Front. Phys. **11**, 1061580 (2023).

[27] J. Kim, G. Pescia, B. Fore, J. Nys, G. Carleo, S. Gandolfi, M. Hjorth-Jensen, and A. Lovato, Neural-network quantum states for ultra-cold Fermi gases, Commun. Phys. **7**, 148 (2024).

[28] C. Adams, G. Carleo, A. Lovato, and N. Rocco, Variational Monte Carlo calculations of $A \leqslant 4$ nuclei with an artificial neural-network correlator ansatz, Phys. Rev. Lett. **127**, 022502 (2021).

[29] A. Gnech, C. Adams, N. Brawand, G. Carleo, A. Lovato, and N. Rocco, Nuclei with up to $A = 6$ nucleons with artificial neural network wave functions, Few-Body Syst. **63**, 7 (2022).

[30] A. Lovato, C. Adams, G. Carleo, and N. Rocco, Hidden-nucleons neural-network quantum states for the nuclear many-body problem, Phys. Rev. Res. **4**, 043178 (2022).

[31] Y. L. Yang and P. W. Zhao, A consistent description of the relativistic effects and three-body interactions in atomic nuclei, Phys. Lett. B **835**, 137587 (2022).

[32] Y. L. Yang and P. W. Zhao, Deep-neural-network approach to solving the *ab initio* nuclear structure problem, Phys. Rev. C **107**, 034320 (2023).

[33] E. Parnes, N. Barnea, G. Carleo, A. Lovato, N. Rocco, and X. Zhang, Nuclear responses with neural-network quantum states, Phys. Rev. Lett. **136**, 032501 (2026).

[34] A. Di Donna, L. Contessi, A. Lovato, and F. Pederiva, Hypernuclei with neural network quantum states, Phys. Rev. Res. **8**, 013160 (2026).

[35] 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).

[36] Y. Nomura, Machine learning quantum states—Extensions to fermion-boson coupled systems and excited-state calculations, J. Phys. Soc. Jpn. **89**, 054706 (2020).

[37] L. L. Viteritti, F. Ferrari, and F. Becca, Accuracy of restricted Boltzmann machines for the one-dimensional J_1–J_2 Heisenberg model, SciPost Phys. **12**, 166 (2022).

[38] 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).

[39] D. Pfau, S. Axelrod, H. Sutterud, I. von Glehn, and J. S. Spencer, Accurate computation of quantum

excited states with neural networks, Science **385**, eadn0137 (2024).
[40] Z. Li, Z. Lu, R. Li, X. Wen, X. Li, L. Wang, J. Chen, and W. Ren, Spin-symmetry-enforced solution of the many-body Schrödinger equation with a deep neural network, Nat. Comput. Sci. **4**, 910 (2024).
[41] P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer, New York, 1980).
[42] S. Inouye, M. R. Andrews, J. Stenger, H. J. Miesner, D. M. Stamper-Kurn, and W. Ketterle, Observation of Feshbach resonances in a Bose–Einstein condensate, Nature (London) **392**, 151 (1998).
[43] C. Chin, R. Grimm, P. Julienne, and E. Tiesinga, Feshbach resonances in ultracold gases, Rev. Mod. Phys. **82**, 1225 (2010).
[44] V. Efimov, Energy levels arising from resonant two-body forces in a three-body system, Phys. Lett. B **33**, 563 (1970).
[45] V. Efimov, Energy levels of three resonantly interacting particles, Nucl. Phys. A **210**, 157 (1973).
[46] T. Kraemer, M. Mark, P. Waldburger, J. G. Danzl, C. Chin, B. Engeser, A. D. Lange, K. Pilch, A. Jaakkola, H. C. Nägerl, and R. Grimm, Evidence for Efimov quantum states in an ultracold gas of Cesium atoms, Nature (London) **440**, 315 (2006).
[47] E. Braaten and H.-W. Hammer, Universality in few-body systems with large scattering length, Phys. Rep. **428**, 259 (2006).
[48] C. H. Greene, P. Giannakeas, and J. Pérez-Ríos, Universal few-body physics and cluster formation, Rev. Mod. Phys. **89**, 035006 (2017).
[49] P. Naidon and S. Endo, Efimov physics: A review, Rep. Prog. Phys. **80**, 056001 (2017).
[50] J. P. D' Incao, Few-body physics in resonantly interacting ultracold quantum gases, J. Phys. B: At. Mol. Opt. Phys. **51**, 043001 (2018).
[51] S. Endo, E. Epelbaum, P. Naidon, Y. Nishida, K. Sekiguchi, and Y. Takahashi, Three-body forces and Efimov physics in nuclei and atoms, Eur. Phys. J. A **61**, 9 (2025).
[52] P. F. Bedaque, H.-W. Hammer, and U. van Kolck, Renormalization of the three-body system with short-range interactions, Phys. Rev. Lett. **82**, 463 (1999).
[53] M. Zaccanti, B. Deissler, C. D'Errico, M. Fattori, M. Jona-Lasinio, S. Müller, G. Roati, M. Inguscio, and G. Modugno, Observation of an Efimov spectrum in an atomic system, Nat. Phys. **5**, 586 (2009).
[54] N. Gross, Z. Shotan, S. Kokkelmans, and L. Khaykovich, Observation of universality in ultracold ^{7}Li three-body recombination, Phys. Rev. Lett. **103**, 163202 (2009).
[55] S. E. Pollack, D. Dries, and R. G. Hulet, Universality in three- and four-body bound states of ultracold atoms, Science **326**, 1683 (2009).
[56] F. Ferlaino, S. Knoop, M. Berninger, W. Harm, J. P. D'Incao, H.-C. Nägerl, and R. Grimm, Evidence for universal four-body states tied to an Efimov trimer, Phys. Rev. Lett. **102**, 140401 (2009).
[57] A. Zenesini, B. Huang, M. Berninger, S. Besler, H.-C. Nägerl, F. Ferlaino, R. Grimm, C. H. Greene, and J. von Stecher, Resonant five-body recombination in an ultracold gas of bosonic atoms, New J. Phys. **15**, 043040 (2013).
[58] R. Pires, J. Ulmanis, S. Häfner, M. Repp, A. Arias, E. D. Kuhnle, and M. Weidemüller, Observation of Efimov resonances in a mixture with extreme mass imbalance, Phys. Rev. Lett. **112**, 250404 (2014).
[59] S.-K. Tung, K. Jiménez-García, J. Johansen, C. V. Parker, and C. Chin, Geometric scaling of Efimov states in a ^{6}Li - ^{133}Cs mixture, Phys. Rev. Lett. **113**, 240402 (2014).
[60] B. D. Esry, C. H. Greene, and J. P. Burke, Recombination of three atoms in the ultracold limit, Phys. Rev. Lett. **83**, 1751 (1999).
[61] J. von Stecher, J. P. D'Incao, and C. H. Greene, Signatures of universal four-body phenomena and their relation to the Efimov effect, Nat. Phys. **5**, 417 (2009).
[62] J. von Stecher, Weakly bound cluster states of Efimov character, J. Phys. B: At. Mol. Opt. Phys. **43**, 101002 (2010).
[63] M. Gattobigio and A. Kievsky, Universality and scaling in the n-body sector of Efimov physics, Phys. Rev. A **90**, 012502 (2014).
[64] P. Naidon, E. Hiyama, and M. Ueda, Universality and the three-body parameter of ^{4}He trimers, Phys. Rev. A **86**, 012502 (2012).
[65] E. Hiyama and M. Kamimura, Universality in Efimov-associated tetramers in ^{4}He, Phys. Rev. A **90**, 052514 (2014).
[66] S. Grebenev, J. P. Toennies, and A. F. Vilesov, Superfluidity within a small helium-4 cluster: The microscopic Andronikashvili experiment, Science **279**, 2083 (1998).
[67] 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).
[68] Y. Yan and D. Blume, Energy and structural properties of n-boson clusters attached to three-body Efimov states: Two-body zero-range interactions and the role of the three-body regulator, Phys. Rev. A **92**, 033626 (2015).
[69] J. Levinsen, P. Massignan, S. Endo, and M. M. Parish, Universality of the unitary Fermi gas: A few-body perspective, J. Phys. B: At. Mol. Opt. Phys. **50**, 072001 (2017).
[70] D. Blume and K. M. Daily, Breakdown of universality for unequal-mass Fermi gases with infinite scattering length, Phys. Rev. Lett. **105**, 170403 (2010).
[71] K. M. Daily and D. Blume, Thermodynamics of the two-component Fermi gas with unequal masses at unitarity, Phys. Rev. A **85**, 013609 (2012).
[72] H. W. Hammer and L. Platter, Efimov states in nuclear and particle physics, Annu. Rev. Nucl. Part. Sci. **60**, 207 (2010).
[73] I. Tanihata, M. Alcorta, D. Bandyopadhyay, R. Bieri, L. Buchmann, B. Davids, N. Galinski, D. Howell, W. Mills, S. Mythili, *et al.*, Measurement of the two-halo neutron transfer reaction ^{1}H(^{11}Li, ^{9}Li) ^{3}H at 3A MeV, Phys. Rev. Lett. **100**, 192502 (2008).
[74] D. Hove, E. Garrido, P. Sarriguren, D. V. Fedorov, H. O. U. Fynbo, A. S. Jensen, and N. T. Zinner, Emergence of clusters: Halos, Efimov states, and experimental signals, Phys. Rev. Lett. **120**, 052502 (2018).
[75] M. Kunitski, S. Zeller, J. Voigtsberger, A. Kalinin, L. P. H. Schmidt, M. Schöffler, A. Czasch, W. Schöllkopf, R. E. Grisenti, T. Jahnke, D. Blume, and R. Dörner, Observation of the Efimov state of the helium trimer, Science **348**, 551 (2015).
[76] Y. Nishida, Y. Kato, and C. D. Batista, Efimov effect in quantum magnets, Nat. Phys. **9**, 93 (2013).
[77] D. S. Petrov, Three-body problem in Fermi gases with short-range interparticle interaction, Phys. Rev. A **67**, 010703(R) (2003).
[78] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).

[79] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv:1412.6980.

[80] G. Pöschl and E. Teller, Bemerkungen zur Quantenmechanik des anharmonischen oszillators, Z. Phys. **83**, 143 (1933).

[81] S. Flügge, *Practical Quantum Mechanics* (Springer, New York, 1974).

[82] L. Platter, C. Ji, and D. R. Phillips, Range corrections to three-body observables near a Feshbach resonance, Phys. Rev. A **79**, 022702 (2009).

[83] F. Schäfer, Y. Haruna, and Y. Takahashi, Observation of Feshbach resonances in an ^{167}Er – ^{6}Li Fermi–Fermi mixture, J. Phys. Soc. Jpn. **92**, 054301 (2023).

[84] K. Xie, X. Li, Y.-Y. Zhou, J.-H. Luo, S. Wang, Y.-Z. Nie, H.-C. Shen, Y.-A. Chen, X.-C. Yao, and J.-W. Pan, Feshbach spectroscopy of ultracold mixtures of ^{6}Li and ^{164}Dy atoms, Phys. Rev. A **111**, 023327 (2025).

[85] K. Oi, P. Naidon, and S. Endo, Universality of Efimov states in highly mass-imbalanced cold-atom mixtures with van der Waals and dipole interactions, Phys. Rev. A **110**, 033305 (2024).

[86] K. Oi and S. Endo, Universal Efimov spectra and fermionic doublets in highly mass-imbalanced cold-atom mixtures with van der Waals and dipole interactions, Phys. Rev. Res. **7**, 033236 (2025).

[87] G. Breit, Separation of angles in the two-electron problem, Phys. Rev. **35**, 569 (1930).

[88] Y. Wang, J. P. D'Incao, and C. H. Greene, Efimov effect for three interacting bosonic dipoles, Phys. Rev. Lett. **106**, 233201 (2011).

[89] Y. Wang, J. P. D'Incao, and C. H. Greene, Universal three-body physics for fermionic dipoles, Phys. Rev. Lett. **107**, 233201 (2011).

[90] T. Shi, H. Wang, and X. Cui, Universal bound states with Bose-Fermi duality in microwave-shielded ultracold molecules, Phys. Rev. Lett. **136**, 043402 (2026).

[91] Y. Ohishi, K. Oi, and S. Endo, Analytical solution of the Schrödinger equation with $1/r^3$ and attractive $1/r^2$ potentials: Universal three-body parameter of mixed-dimensional Efimov states, Phys. Rev. Res. **8**, 023101 (2026).

[92] See Supplemental Material at http://link.aps.org/supplemental/10.1103/31mv-7tz7 for numerical codes written in PyTorch framework. These codes include calculations of (1) the ground states of identical bosons, (2) the corresponding first excited states, (3) the ground state of two identical fermions and one particle, and (4) the corresponding first excited state.