Nuclear Science and Techniques (2025) 36:196
https://doi.org/10.1007/s41365-025-01775-4

# Optimizing basis wave functions in the generator coordinate method for microscopic cluster models (I)

Yi-Fan Liu[1] · Bo Zhou[1,2] · Yu-Gang Ma[1,2]

Received: 8 January 2025 / Revised: 31 January 2025 / Accepted: 19 February 2025 / Published online: 7 August 2025


**Abstract**
We employed random distributions and gradient descent methods for the Generator Coordinate Method (GCM) to identify effective basis wave functions, taking halo nuclei $^6$He and $^6$Li as examples. By comparing the ground state ($0^+$) energy of $^6$He and the excited state ($0^+$) energy of $^6$Li calculated with various random distributions and manually selected generation coordinates, we found that the heavy tail characteristic of the logistic distribution better describes the features of the halo nuclei. Subsequently, the Adam algorithm from machine learning was applied to optimize the basis wave functions, indicating that a limited number of basis wave functions can approximate the converged values. These results offer some empirical insights for selecting basis wave functions and contribute to the broader application of machine learning methods in predicting effective basis wave functions.

**Keywords** Generator Coordinate Method · Effective basis wave functions · Nuclear cluster model · Machine learning · Halo nuclei

## 1 Introduction

Clustering is a universal phenomenon observed in various systems, ranging from clusters of galaxies to clusters of nuclei [1–3]. In nuclear physics, clustering is one of the most important features in light nuclei [4–9]. Since the development of the $\alpha$ cluster model, light nuclei have been studied from the perspective of cluster features for more than half a century [10–14]. Various nuclear theories have been proposed to investigate nuclear clustering [15–19]. The three traditional nuclear cluster models are the Resonance Group Method (RGM) [20–22], Generator Coordinate Method (GCM) [23–26], and Orthogonality Condition Model (OCM) [27]. Some developed models, such as the Antisymmetrized Molecular Dynamics (AMD) model [28–30] and the Tohsaki-Horiuchi-Schuck-Röpke (THSR) model [31], have also been introduced in recent years.

The Generator Coordinate Method (GCM) was first introduced by Hill and Wheeler [23] in 1953 in the context of nuclear fission. Subsequently, Griffin and Wheeler extended this method of a general many-body tool [24]. The principle of the GCM is to express nuclear state wave functions as superpositions of non-orthogonal basis functions such as Slater determinants [32]. Because of the flexibility in selecting the basis functions or generator coordinates, GCM offers a general method for addressing many-body problems in nuclear cluster physics and other fields [33–37].

The GCM requires the superposition of different types of basis wave functions, thereby demonstrating their flexibility. However, selecting basis states is a crucial issue in some cases. The choice of collective coordinates often relies on empirical and phenomenological methods, which increase the

This work was supported by the National Key R&D Program of China (No. 2023YFA1606701), the National Natural Science Foundation of China (Nos. 12175042, 11890710, 11890714, 12047514, 12147101, and 12347106), Guangdong Major Project of Basic and Applied Basic Research (No. 2020B0301030008), and China National Key R&D Program (No. 2022YFA1602402).

✉ Bo Zhou
zhou_bo@fudan.edu.cn

✉ Yu-Gang Ma
mayugang@fudan.edu.cn

[1] Key Laboratory of Nuclear Physics and Ion-beam Application (MOE), Institute of Modern Physics, Fudan University, Shanghai 200433, China

[2] Shanghai Research Center for Theoretical Nuclear Physics, NSFC and Fudan University, Shanghai 200438, China

complexity and computational time for many-body cluster systems. This issue is especially pronounced when applying the GCM to the structure of halo nuclei such as $^6$He, a well-known Borromean nucleus. It comprises a loosely bound and spatially extended three-body system, typically including the $\alpha$ core surrounded by two weakly bound neutrons $\alpha + \mathrm{n} + \mathrm{n}$ [38–40]. Using more efficient basis states in the GCM to describe such three-body gas-like systems accurately is an important issue [41].

In recent years, there have been many theoretical studies [42–44] exploring how to select effective basis states for the GCM. For example, Suzuki and Varga introduced stochastic sampling in the few-body model [45]; Suhara and Kanada-En'yo have proposed the $\beta - \gamma$ constrained selection of Slater determinants in nuclear cluster model [46]; Additionally, Fukuoka et al. developed the imaginary-time evolution method in the mean-field model [47]; And Takatoshi et al. refined the Bloch-Brink $\alpha$ cluster model with the stochastic sampling method [48]. Owing to the powerful data processing capabilities of machine learning algorithms (ML), they have been widely employed in addressing different nuclear physics issues, including nuclear mass systematics [49–51], radii prediction [52], decay descriptions [53], many-body problems [54, 55], and nuclear structure [56]. It has attempted to identify hidden laws from a large amount of historical data and use them for prediction or classification.

In this study, by choosing the di-neutron halo nucleus of $^6$He ($\alpha + \mathrm{n} + \mathrm{n}$) and the proton–neutron halo of $^6$Li ($\alpha + \mathrm{n} + \mathrm{p}$) systems [57], we studied the effective basis problems in GCM, using global optimization and local gradient descent methods. First, the empirical law of the effective basis wave function distribution was summarized by comparing various random distributions. Subsequently, the Adam method [58] was used to provide a better standard of basis wave functions.

This paper is organized in the following way. Section 2 briefly reviews the framework of the wave function and the Generator Coordinate Method. In Sect. 3, the numerical results of optimization and discussions are provided. Finally, a summary is provided in Sect. 4.

## 2 Theoretical framework

The general ansatz of the Generator Coordinate Method [24] can be expressed,

$$|\Psi_{\mathrm{GCM}}\rangle = \int \mathrm{d}\boldsymbol{q}|\Phi(\boldsymbol{q})\rangle f(\boldsymbol{q}). \tag{1}$$

where $\boldsymbol{q} = (q_1, q_2, ..., q_n)$ denotes a series of generator coordinates. The $f(\boldsymbol{q})$ is the weight function. The trial wave function $|\Phi(\boldsymbol{q})\rangle$ is important, because it should be based on the specific motion patterns of the physical system. The GCM approach offers the benefit of obtaining the ground states and various categories of excited states that are described by the selected generator coordinates.

In nuclear cluster physics, the Brink wave function [59] is typically used as the basis wave function for GCM calculations. Taking $^6$He as an example, the Brink wave function with an $\alpha + \mathrm{n} + \mathrm{n}$ cluster configuration can be written as

$$\begin{aligned}
&\Phi(\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_\alpha) = \mathcal{A}\{\Phi_{\mathrm{n}}(\boldsymbol{R}_1)\Phi_{\mathrm{n}}(\boldsymbol{R}_2)\Phi_\alpha(\boldsymbol{R}_\alpha)\} \\
&\Phi_\alpha(\boldsymbol{R}_\alpha) = \mathcal{A}\left\{\prod_{i=3}^{6}\phi(\boldsymbol{R}, \boldsymbol{r}_i)\chi_i\tau_i\right\} \\
&\Phi_n(\boldsymbol{R}) = \phi(\boldsymbol{R}, \boldsymbol{r}_i)\chi_i\tau_i \\
&\phi(\boldsymbol{R}, \boldsymbol{r}_i) = \left(\frac{1}{\pi b^2}\right)^{3/4} e^{-\frac{(r_i-R)^2}{2b^2}}.
\end{aligned} \tag{2}$$

Here, the wave function $\Phi_\alpha$ represents the $\alpha$-cluster with a configuration of $(0s)^4$, and the valence neutron wave function is denoted as $\Phi_{\mathrm{n}}$. $\boldsymbol{R}_1, \boldsymbol{R}_2$ and $\boldsymbol{R}_\alpha$ represent the generator coordinates of $\alpha$ particles and two neutrons, abbreviated as $\{\boldsymbol{R}\} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_\alpha\}$. $\phi(\boldsymbol{R}, \boldsymbol{r}_i)\chi_i\tau_i$ describes the $i$-th single-particle wave function, with $\phi(\boldsymbol{R}, \boldsymbol{r}_i)$ specifying the spatial wave function. The spin and isospin of each nucleon are denoted by $\chi_i$ and $\tau_i$, respectively. The spins of the two valence neutrons were set up and down, respectively. The harmonic oscillator parameter $b = \sqrt{1/(2\nu)} = 1.46$ fm to avoid spurious center-of-mass problems in this work, which is identical to that used in Ref. [60, 61]. The microscopic cluster wave function of $^6$Li($\alpha + \mathrm{n} + \mathrm{p}$) can be constructed similarly.

Within the GCM framework, the final wave function of $^6$He can be obtained by superposing various configurations of $\alpha + \mathrm{n} + \mathrm{n}$.

$$\Psi_M^{J\pi} = \sum_{\{\boldsymbol{R}\}K} f_{\{\boldsymbol{R}\}K} P^J_{MK} P^\pi \Phi(\{\boldsymbol{R}\}), \tag{3}$$

where $P^J_{MK}$ and $P^\pi$ denote the angular-momentum and parity projector, respectively. For convenience, we write $P^J_{MK}P^\pi\Phi(\{\boldsymbol{R}\}) = \Phi^{J\pi}_{\mathrm{MK}}(\{\boldsymbol{R}\})$. The coefficients $f_{\{\boldsymbol{R}\}K}$ can then be calculated using the Hill–Wheeler equation [32]

$$\begin{aligned}
\sum_{\{\boldsymbol{R}'\}K'} f_{\{\boldsymbol{R}'\}K'} \Big[ &\left\langle \Phi^{J\pi}_{MK}(\{\boldsymbol{R}\}) \middle| \hat{H} \middle| \Phi^{J\pi}_{MK'}(\{\boldsymbol{R}'\}) \right\rangle \\
&- E\left\langle \Phi^{J\pi}_{MK}(\{\boldsymbol{R}\}) \mid \Phi^{J\pi}_{MK'}(\{\boldsymbol{R}'\}) \right\rangle \Big] = 0.
\end{aligned} \tag{4}$$

By changing the values in the generated coordinate set $\{\boldsymbol{R}\}$, various basis wave functions can be obtained. Note that as long as the number of superposed basis wave functions is sufficiently large, the final wave function is highly accurate. However, this approach results in a significant computational time. Fortunately, not all the basis wave functions are equally important. Therefore, selecting effective basis wave

functions can achieve the same effect as superposing a large number of them, but with fewer effective wave functions. The objective of this study is to select effective basis wave functions characterized by the generated coordinate set $\{\boldsymbol{R}\}$ from both global and local perspectives.

The Hamiltonian for $^6$He and $^6$Li three-body systems can be written as:

$$\hat{H} = \sum_{i=1}^{A} \hat{t}_i - \hat{T}_{\text{c.m.}} + \sum_{i<j}^{A} \hat{v}_{\text{N}} + \sum_{i<j}^{A} \hat{v}_{\text{C}} + \sum_{i<j}^{A} \hat{v}_{\text{LS}}, \tag{5}$$

where $t_i$ represents the kinetic energy of the individual nucleons, and the center of mass is denoted by $T_{\text{c.m.}}$. $v_{\text{N}}$, $v_{\text{C}}$, and $v_{\text{LS}}$ denote the effective nucleon–nucleon interaction, the Coulomb interaction and spin-orbit interaction, respectively. The Volkov No.2 interaction [62] was employed for the nucleon–nucleon interaction. The expression is given as:

$$\begin{aligned}\hat{v}_{\text{N}} =& \left(W - M\hat{P}^{\sigma}\hat{P}^{\tau} + B\hat{P}^{\sigma} - H\hat{P}^{\tau}\right) \\ &\times \left[V_1 \exp\left(-r^2/c_1^2\right) + V_2 \exp\left(-r^2/c_2^2\right)\right].\end{aligned} \tag{6}$$

The parameters were set as follows: $W = 0.4$, $M = 0.6$. For $^6$He, $B = H = 0.125$ [10]; For $^6$Li, $B = H = 0.08$ [63]. Regarding the Gaussian terms, the values are $V_1 = -60.65$ MeV, $V_2 = 61.14$ MeV, $c_1 = 1.80$ fm, and $c_2 = 1.01$ fm.

For the spin–orbit interaction, the G3RS potential [64, 65] is adopted,

$$\hat{v}_{\text{LS}} = V_0\left(e^{d_1 r^2} - e^{d_2 r^2}\right)\hat{P}_{31}\hat{L}\cdot\hat{S}. \tag{7}$$

The strength parameter $V_0$ was fixed at 2000 MeV. The Gaussian parameters $d_1$ and $d_2$ are configured to 5.0 fm$^{-2}$ and 2.778 fm$^{-2}$, respectively.

## 3 Results and discussion

In this section, we take the di-neutron halo nucleus $^6$He and proton–neutron halo nucleus $^6$Li as examples, each conceptualized as three-cluster structures $\alpha + \text{n} + \text{n}$ and $\alpha + \text{n} + \text{p}$, respectively. First, we summarize empirical laws using random distribution methods from a global perspective. Subsequently, at the local level, the Adam algorithm from machine learning is introduced to optimize the generation coordinates using the gradient descent theory.

### 3.1 Searching for specific distributions leading to effective basis wave functions

In the GCM, the mesh points for the generator coordinates were not predetermined. Although it is theoretically feasible to obtain exact solutions by enumerating a large number of wave functions with different configurations, this approach is computationally impractical. Instead, we hypothesize that the effective basis wave functions may follow specific distributions. To test this hypothesis, we generated coordinate sets $\{\boldsymbol{R}\}$ using various random distributions and applied them to calculate the ground state energy of the di-neutron halo nucleus $^6$He. By analyzing the relationship between the number of superimposed basis wave functions and the resulting ground state energies, we aim to derive global empirical rules that can enhance the efficiency of the GCM.

As a benchmark test, we first calculated the ground state of the di-neutron halo $^6$He nucleus using traditional mesh points for the Brink wave functions. As shown in Fig. 1, different sets of coordinates can be generated by adjusting the relative distances $L_1$, $L_2$, and angle $\theta$ relative to the $x$-axis. Where $L_1$ is the distance between the $\alpha$ particle and neutron, and $L_2$ is the distance between the other neutron and the center-of-mass of the $\alpha$ particle and neutron. Subsequently, various configurations with different sets $(L_1, L_2, \theta)$ were superposed. The mesh points for $L_1$ were established at intervals of 0.35 fm, resulting in 14 points, whereas for $L_2$, the intervals were set at 0.5 fm, accumulating 18 points. The values for angle $\theta$ were sequentially determined to be 0°, 30°, 60°, and 90°. By employing this traditional method, 1008 basis wave functions were generated. Subsequently, the ground state energy was computed after diagonalization. Due to large overlap between some wave functions or their lack of physical significance, 539 out of the initial 1008 basis wave functions were removed during the calculations. This outcome is shown in the last column of Table 1, along with other results recorded in the same table. Furthermore, Fig. 2 displays how the energy varies with the superposition of wave functions.

Another common method for selecting mesh points is to distribute them within a spherical shell structure. The spherical structure of the latter is depicted as a three-dimensional, multi-layered spherical shell. The initial radius of the spherical shell was set at 1 fm, with spacing between adjacent layers at 1 fm, producing a total of five layers from the inside

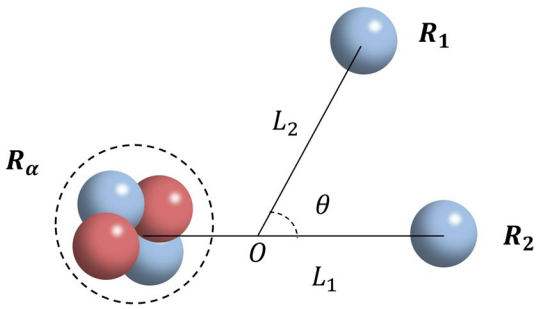


**Fig. 1** (Color online) This schematic figure illustrates the Brink wave function of $^6$He ($\alpha + \text{n} + \text{n}$). Blue spheres represent neutrons, while red spheres denote protons. The $\alpha$ cluster and the two neutrons are aligned in the same plane

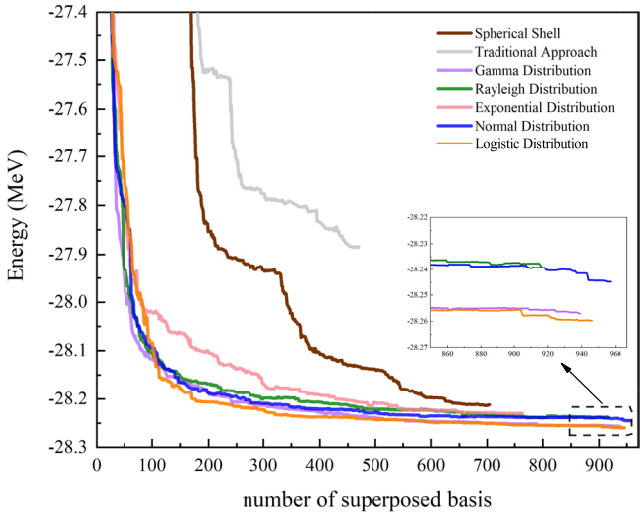


**Fig. 2** (Color online) Energy convergence of the ground state ($0^+$) of $^6$He using various methods to generate basis wave functions. The brown and gray lines represent the results from the spherical shell and traditional methods, respectively. The curves in other colors represent the calculation result of the generated coordinates generated by the random distribution

out. The Marsaglia algorithm [66] was employed to ensure a uniform distribution of points on each spherical shell. For each layer, 200 random three-dimensional coordinate values were generated. Following a procedure similar to the earlier one, the coordinates of the 1000 points generated are designated as $\mathbf{R}_{1i}(x_{1i}, y_{1i}, z_{1i})$. This randomization process was repeated three times to obtain the generation coordinates for the $\alpha$ particles and two neutrons. By superposing these 1000 basis wave functions, the ground state of $^6$He was calculated. Detailed information regarding the results is presented in Fig. 2 and Table 1.

In addition to the above common methods, we also introduce other random distribution functions for generating mesh points, such as Gamma distribution, Uniform distribution, Chi-square distribution, Logistic distribution, and Normal distribution. Considering the halo feature of $^6$He, which is characterized by a diffuse density distribution around the nucleus, specific parameters for various random distribution functions have been selectively determined to assess their impact on the tail region. Although no quantitative relationships for parameter values under different distributions have been specified, we fortuitously discovered that the results remained relatively stable within a reasonable range of parameter values after experimenting with various settings. The probability density functions and specific parameter values for random distributions are listed in Table 1. Furthermore, GCM calculations are performed within the center-of-mass coordinate system; consequently, the results depend solely on the relative distances between the clusters. Thus, in the Normal and Logistic distributions, the results are influenced only by the standard deviation $\sigma$, independent of the mean $\mu$. For computational convenience, $\mu$ is set to zero in these cases. Each random distribution is utilized to generate 1000 basis wave functions for the GCM calculations.

Figure 2 shows the energy convergence of the ground state using the various methods presented here to generate the basis wave functions. As one can see, the convergence rate of the energy variation curves corresponding to random distributions is significantly faster than that derived from manually configured structures. Table 1 provided qualitative insights. The ground state energies obtained using the traditional and spherical shell methods are −27.884 MeV and −28.212 MeV, respectively, both of which are obviously higher than those from all random distributions.

From Fig. 2 and Table 1, it is interesting to see that the Logistic distributions outperform other random distributions, with the $0^+$ state energies calculated at −28.260 MeV. Note that the Gamma distribution is also effective, although it is determined by two parameters $(\alpha, \beta)$, which provide a wider range of adjustments. Single-parameter Normal and Logistic distributions offer significant advantages for

**Table 1** Comparison of the different methods to generate basis wave functions. The calculated ground state energies of $^6$He (MeV) and the number of superposed basis are listed

| Distribution model | Probability density function | Parameters | $E$ (MeV) | Superposed basis |
|---|---|---|---|---|
| Spherical shell | – | – | −28.212 | 704 |
| Traditional approach | – | – | −27.884 | 469 |
| Gamma distribution | $f(x;\alpha,\beta)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}$ | $\alpha=2.3,\beta=1.5$ | −28.257 | 939 |
| Rayleigh distribution | $f(x;\sigma)=\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}$ | $\sigma=3.0$ | −28.239 | 916 |
| Exponential distribution | $f(x;\lambda)=\lambda e^{-\lambda x}$ | $\lambda=0.5$ | −28.229 | 764 |
| Normal distribution | $f(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | $(\mu=0),\sigma=2.3$ | −28.245 | 957 |
| Logistic distribution | $f(x;\mu,\gamma)=\frac{e^{-(x-\mu)/\gamma}}{\gamma\left(1+e^{-(x-\mu)/\gamma}\right)^2}$ | $(\mu=0),\gamma=1.2$ | −28.260 | 946 |

practical calculations. This superiority is attributed to the assumption of independence in the Normal distribution, which mirrors the relative spatial near independence of the clusters in the nuclei. Furthermore, the thicker asymptotic tails of the probability density function in the Normal distribution closely corresponded to the halo characteristics of $^6$He. Similarly, although the Logistic distribution resembles the Normal distribution, it features notably heavier tails. This characteristic better captures the extended features of the $^6$He halo nucleus structure, thereby encompassing a broader distribution of the effective basis states.

To substantiate this conclusion further, we compared the ground state energy by maintaining approximate equality between the standard deviations of the Logistic and Normal distributions according to the theoretical formula $\gamma = (\sqrt{3}\sigma)/\pi$. In this case, the two distributions have similar shapes but slightly different tail thicknesses. We selected four sets of generation coordinates with varying parameters for comparison, where each set generates 1000 basic wave functions. Specifically, the parameters for the Logistic distribution were set at $\gamma = 1.2, 1.3, 1.4, 1.5$ and for the Normal distribution at $\sigma = 1.8, 2.0, 2.3, 2.5$. The calculated ground state ($0^+$) energies of $^6$He are presented in Table 2, and Fig. 3 illustrates the energy convergences with an increasing number of basis. From Table 2, it can be seen that among the computations employing Normal distributions, the best performance is achieved with a standard deviation $\sigma$ of 2.3, where the ground state energy converges to −28.245 MeV. Conversely, for $\sigma$ values of 1.8, 2.0, and 2.5, the energies converge to −28.224 MeV, −28.222 MeV, and −28.231 MeV, respectively. On the other hand, the Logistic distribution clearly outperforms the Normal distribution, with ground state energies uniformly converging around −28.26 MeV. Notably, even at its poorest performance with parameter $\gamma = 1.2$, the energy convergence reaches −28.260 MeV, surpassing the mean performance of the Normal distribution groups. Additionally, Fig. 3 illustrates that the rate of energy convergence for the Logistic distribution group is substantially faster than that for the Normal distribution group, thereby exhibiting greater robustness. This indicates that the Logistic distribution includes a more effective basis, which could be because of its heavy tail part.

**Table 2** The numerical results of the ground state ($0^+$) of $^6$He, from the calculations of Logistic distribution and Normal distribution with different parameters

| Distribution | Parameters | $E$ (MeV) |
|---|---|---|
| Logistic distribution | $\gamma = 1.2$ | −28.260 |
| | $\gamma = 1.3$ | −28.268 |
| | $\gamma = 1.4$ | −28.262 |
| | $\gamma = 1.5$ | −28.266 |
| Normal distribution | $\sigma = 1.8$ | −28.224 |
| | $\sigma = 2.0$ | −28.222 |
| | $\sigma = 2.3$ | −28.245 |
| | $\sigma = 2.5$ | −28.231 |

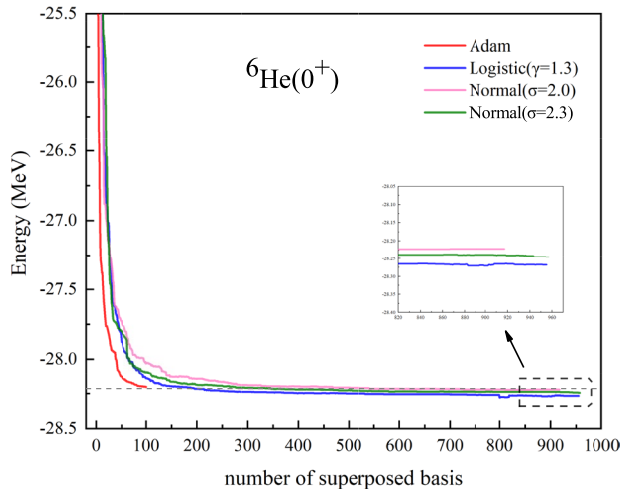


**Fig. 3** (Color online) Energy convergence of the ground state ($0^+$) of $^6$He. To compare the results of Normal distribution and Logistic distribution, 1000 ground state wave functions are generated using four sets of different parameters. In addition, the red line denotes the results of Adam optimization

It is worth mentioning that the ground state ($1^+$) of $^6$Li converges rapidly, requiring only a minimal number of basis wave functions. Calculations using various parameters for the Logistic and Normal distributions also quickly converged to approximately −30.02 MeV, indicating that there is no need for further optimization of the basis wave functions. Therefore, further discussion of the $^6$Li ground state is omitted in this work.

To confirm the aforementioned conclusions and ascertain the universality of the Logistic distribution in halo nuclear structures, we studied the excited states ($0^+$) of $^6$Li and generated three sets of Logistic and Normal distributions under conditions of similar standard deviations for comparative analysis. Parameters for the Logistic distribution were set at $\gamma$ =1.0, 1.3, and 1.5, whereas those for the Normal distribution were set at $\sigma = 1.8$, 2.0, and 2.3. Considering the proton–neutron halo structure of $^6$Li, whose excited state energy converges more readily than the ground state energy of $^6$He, we generated 400 basis wave functions for each parameter set to perform the GCM calculations. The results are shown in Fig. 4 and Table 3. It is gratifying to observe that compared with the Normal distribution, the Logistic distribution still performs well in each group. According to the excitation energy of the $0^+$ state in Table 3, it can be observed that the Logistic distribution parameters $\gamma$ are set at 1.0, 1.3, and 1.5, yielding

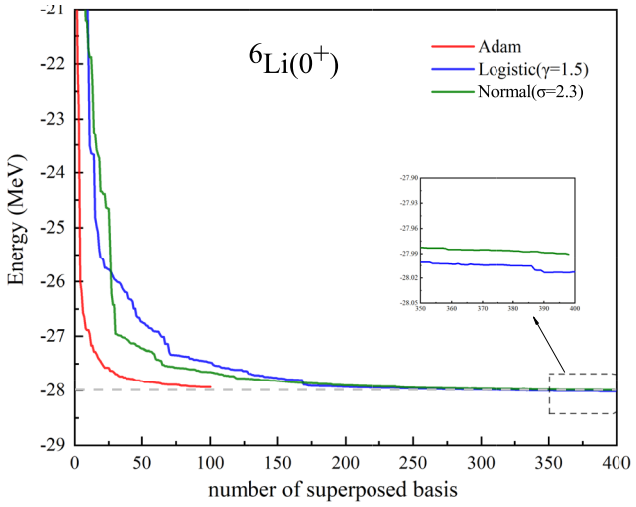


**Fig. 4** (Color online) Energy convergence of the excited state ($0^+$) of $^6$Li. To compare the results of Normal distribution and Logistic distribution, 400 wave functions are generated using three sets of different parameters. The red line represents the results of Adam optimization

**Table 3** The numerical results of the excited state ($0^+$) of $^6$Li, from the calculations of Logistic distribution and Normal distribution with different parameters

| Distribution | Parameters | $E$ (MeV) |
|---|---|---|
| Logistic distribution | $\gamma = 1.0$ | −27.945 |
| | $\gamma = 1.3$ | −27.999 |
| | $\gamma = 1.5$ | −28.013 |
| Normal distribution | $\sigma = 1.8$ | −27.933 |
| | $\sigma = 2.0$ | −27.969 |
| | $\sigma = 2.3$ | −27.991 |

convergence values of energy at −27.945 MeV, −27.999 MeV, and −28.013 MeV, respectively. These values slightly surpass those derived from the Normal distribution parameter $\sigma$ set at 1.8, 2.0, and 2.3, which resulted in energy convergences of −27.933 MeV, −27.969 MeV, and −27.991 MeV. As shown in Fig. 4, the energy gradually converges as the number of basis wave functions increases, with the Logistic distribution exhibiting a slightly faster rate of convergence than the Normal distribution. These results indicate that the thick-tail characteristic of the Logistic distribution is not only suitable for calculating the ground state ($0^+$) energy of $^6$He with a di-neutron halo, but also applicable to the excited state ($0^+$) of $^6$Li with a proton–neutron halo. Furthermore, the Logistic distribution encompasses a broader range of effective basis wave functions, thereby providing crucial empirical insights for the subsequent selection of effective basis wave functions. It is important to analyze the underlying mechanisms behind this distribution.

### 3.2 Adam optimization based on gradient descent principle

In GCM calculations, determining an appropriate distribution of optimized basis wave functions is of paramount importance for practical computations. However, studying the optimal basis wave functions mathematically at a local level is indispensable.

As shown in Eq. (3), the eigen energy $E$ can be considered as a multivariable function with a set of $\{\boldsymbol{R}\}$ as its independent variables. Thus, solving for the energy using the variational principle is analogous to finding the minimum of a multivariable function $f(\{\boldsymbol{R}\}_1, \{\boldsymbol{R}\}_2, \ldots, \{\boldsymbol{R}\}_i)$. The computational procedure for optimizing the superposition of 100 basis wave functions is as follows: To begin this process, a random set of generated coordinates $\{\{\boldsymbol{R}\}_1, \{\boldsymbol{R}\}_2, \ldots, \{\boldsymbol{R}\}_i\}$ is created as the starting position. For the first basis wave function $\Phi(\{\boldsymbol{R}\}_1)$, the energy gradient must be calculated to adjust the coordinates $\{\boldsymbol{R}\}_1$ in the gradient direction. When a local minimum was reached, this position was recorded as the effective basis wave function location $\{\boldsymbol{R}\}_1^{\text{opt}}$. Next, for the second basis wave function $\Phi(\{\boldsymbol{R}\}_2)$, we superimpose it onto the previously optimized $\{\boldsymbol{R}\}_1^{\text{opt}}$. This results in a simplified expression for the wave function as $\Phi(\{\boldsymbol{R}\}_1^{\text{opt}}) + \Phi(\{\boldsymbol{R}\}_2)$. Then, only coordinates $\{\boldsymbol{R}\}_2$ are adjusted in the direction of the gradient, while coordinates $\Phi(\{\boldsymbol{R}\}_1^{\text{opt}})$ are held constant. Once the local minimum is achieved, the current position is designated as the effective basis wave function position, $\Phi(\{\boldsymbol{R}\}_2^{\text{opt}})$. For the third basis wave function $\Phi(\{\boldsymbol{R}\}_3)$, the same procedure was followed. The adjustment process focuses solely on coordinates $\{\boldsymbol{R}\}_3$ in the direction of the gradient, keeping $\{\boldsymbol{R}\}_1^{\text{opt}}$ and $\{\boldsymbol{R}\}_2^{\text{opt}}$ constant. This method was repeated for each subsequent wave function, ensuring that the optimized coordinates remained fixed, while the new coordinates underwent gradient descent. This iterative approach continues until all basis wave functions are optimized. In principle, a set of optimal generating coordinates for the basis wave functions can be obtained with sufficient computational precision.

We used the Adam algorithm for gradient descent [58]. The Adam algorithm was chosen for its efficiency in optimizing the basis wave functions, particularly for handling complex and flat energy surfaces. By adapting learning rates and leveraging historical gradients, Adam outperformed traditional methods, reducing computational costs and improving the convergence speed. This makes it well suited for challenging calculations, such as those involving halo nuclei.

Given the gradual slowing of energy convergence with increasing number of basis wave functions, distinct treatments were applied to the initial and terminal functions. In the early superposition phase, owing to the rapid decrease in energy, the Adam algorithm's learning rate $\alpha$ was set at 0.3 with a maximum iteration count of 12 and an allowance for four oscillations to prevent missing minimal value points. When the fifth wave function was reached, as the rate of energy decline slowed, the learning rate was adjusted to 0.7 with the maximum iterations increasing to 40, and oscillations allowed up to eight to minimize time consumption.

Although Sect. 3.1 shows the Logistic distribution outperforming Normal distributions and other models, we used the Normal distribution with $\sigma = 2.3$ to generate 100 initial coordinates $\{\boldsymbol{R}\}$ for a clearer comparative analysis. The optimization results for the ground state ($0^+$) of $^6$He, as shown in Fig. 3. Compared to −28.064 MeV with unoptimized basis wave functions, the energy significantly decreases to −28.203 MeV after Adam optimization. Even compared to the well-performing Normal distribution, the superposition of 100 optimized basis wave functions through the Adam method equates to the effects of adding 200 or even 500 basis wave functions in the Normal distribution. This outcome not only confirms the feasibility and effectiveness of gradient descent optimization using the Adam method but also demonstrates that a small number of basis wave functions can achieve the effects of superposing multiple wave functions, and that Adam optimization includes a large number of effective basis wave functions.

To further validate this conclusion, the Adam algorithm was applied to the calculation of the excited state ($0^+$) energy of $^6$Li, and the results are depicted in Fig. 4. Compared to −27.656 MeV with non-optimized basis wave functions, the energy decreased significantly to −27.937 MeV after Adam optimization. Compared with other parameter settings in the Normal distribution, the superposition effect of 100 optimized basis wave functions is equivalent to adding 200−300 ones in the Normal distribution. This suggests that there is still room for improvement in determining better distributions for the mesh points in the GCM. It is worth noting that owing to the differing nuclear structures of various nuclei, the optimized basis functions derived by the Adam algorithm for one nucleus in principle cannot be directly applied to another.

## 4 Summary and outlook

To conclude, we investigated the effective basis wave function distributions for the di-neutron halo nucleus $^6$He and proton–neutron halo nucleus $^6$Li using the Generator Coordinate Method. From a global perspective, our comparative analysis of various random distributions against manually configured models revealed that the Normal distribution performed significantly better than the other distributions, except for the Logistic distribution. This superior performance is attributed to the independence assumption of the Normal distribution, which mirrors the relative spatial independence of the nuclei, and the heavy tail of its probability density function, which resembles the diffuse characteristics typical of halo nuclei. Interestingly, the Logistic distribution, with its probability density curve similar to that of the Normal distribution but with a more pronounced tail, not only retained the advantages of the Normal distribution but also more accurately represented halo nuclear structures. In both the ground state ($0^+$) of $^6$He and the excited state ($0^+$) of $^6$Li, the Logistic distribution yielded better results than the Normal distribution, suggesting that it encompasses a broader range of effective basis wave functions and reduces the number of necessary wave function overlays for halo structures.

Furthermore, from a local viewpoint, we analogized the optimization of basis wave functions to solve a multivariate extremum problem and employed the Adam algorithm to optimize 100 basis wave functions, achieving considerable outcomes. The results for both the ground state ($0^+$) of $^6$He and the excited state ($0^+$) of $^6$Li indicated that a few optimized basis wave functions could achieve the effects of multiple wave functions overlays, thus validating the feasibility and universality of the gradient descent method. This not only provides a benchmark for selecting effective basis wave functions but also makes a solid foundation for future applications of machine learning methods to comprehensively predict the coordinates of effective basis wave functions. Although the Adam algorithm currently requires extensive runtime due to computational constraints, future research will focus on optimizing it to enhance computational efficiency.

**Author Contributions** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yi-Fan Liu and Bo Zhou. The first draft of the manuscript was written by Yi-Fan Liu and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## Declarations

**Conflict of interest** Yu-Gang Ma is the editor-in-chief for Nuclear Science and Techniques and was not involved in the editorial review, or the decision to publish this article. All authors declare that there are no conflict of interest.

## References


1. M. Freer, H. Horiuchi, Y. Kanada-En'yo et al., Microscopic clustering in light nuclei. Rev. Mod. Phys. **90**, 035004 (2018). https://doi.org/10.1103/RevModPhys.90.035004
2. C.Z. Shi, Y.G. Ma, $\alpha$-clustering effect on flows of direct photons in heavy-ion collisions. Nucl. Sci. Tech. **32**, 66 (2021). https://doi.org/10.1007/s41365-021-00897-9
3. Y. Ye, X. Yang, H. Sakurai et al., Physics of exotic nuclei. Nat. Rev. Phys. (2024). https://doi.org/10.1038/s42254-024-00782-5
4. I. Tanihata, H. Toki, T. Kajino, *Handbook of Nuclear Physics*, (Springer Nature, 2023). https://doi.org/10.1007/978-981-15-8818-1
5. W.B. He, Y.G. Ma, X.G. Cao et al., Giant dipole resonance as a fingerprint of $\alpha$ clustering configurations in $^{12}$C and $^{16}$O. Phys. Rev. Lett. **113**, 032506 (2014). https://doi.org/10.1103/PhysRevLett.113.032506
6. B.S. Huang, Y.G. Ma, W.B. He, Alpha-clustering effects on $^{16}$O ($\gamma$, np) $^{14}$N in the quasi-deuteron region. Euro. Phys. J. A **53**, 119 (2017). https://doi.org/10.1140/epja/i2017-12300-0
7. Y.G. Ma, Effects of $\alpha$-clustering structure on nuclear reaction and relativistic heavy-ion collisions. Nucl. Tech. (in Chinese) **46**, 080001 (2023). https://doi.org/10.11889/j.0253-3219.2023.hjs.46.080001
8. Y. Liu, Y.L. Ye, Nuclear clustering in light neutron-rich nuclei. Nucl. Sci. Tech. **29**, 184 (2018). https://doi.org/10.1007/s41365-018-0522-x
9. S. Huang, Z. Yang, Neutron clusters in nuclear systems. Front. Phys. **11**, 1233175 (2023). https://doi.org/10.3389/fphy.2023.1233175
10. Q. Zhao, B. Zhou, M. Kimura et al., Microscopic calculations of $^{6}$He and $^{6}$Li with real-time evolution method. Euro. Phys. J. A **58**, 25 (2022). https://doi.org/10.1140/epja/s10050-021-00648-9
11. Y.Z. Wang, S. Zhang, Y.G. Ma, System dependence of away-side broadening and $\alpha$-clustering light nuclei structure effect in dihadron azimuthal correlations. Phys. Lett. B **831**, 137198 (2022). https://doi.org/10.1016/j.physletb.2022.137198
12. S.M. Wang, J.C. Pei, F.R. Xu, Spectroscopic calculations of cluster nuclei above double shell closures with a new local potential. Phys. Rev. C **87**, 014311 (2013). https://doi.org/10.1103/PhysRevC.87.014311
13. Z.W. Xu, S. Zhang, Y.G. Ma et al., Influence of $\alpha$-clustering nuclear structure on the rotating collision system. Nucl. Sci. Tech. **29**, 186 (2018). https://doi.org/10.1007/s41365-018-0523-9
14. D.Y. Tao, B. Zhou, The reduced-width amplitude in nuclear cluster physics. (2024). arxiv:2412.20928
15. B. Zhou, A. Tohsaki, H. Horiuchi et al., Breathing-like excited state of the hoyle state in $^{12}$C. Phys. Rev. C **94**, 044319 (2016). https://doi.org/10.1103/PhysRevC.94.044319
16. L. Wang, J. Liu, R. Wang et al., Global analysis of nuclear cluster structure from the elastic and inclusive electron scattering. Phys. Rev. C **103**, 054307 (2021). https://doi.org/10.1103/PhysRevC.103.054307
17. B. Zhou, Y. Funaki, H. Horiuchi et al., Nonlocalized clustering: a new concept in nuclear cluster structure physics. Phys. Rev. Lett. **110**, 262501 (2013). https://doi.org/10.1103/PhysRevLett.110.262501
18. B. Zhou, Y. Funaki, H. Horiuchi et al., Nonlocalized clustering and evolution of cluster structure in nuclei. Front. Phys. **15**, 1–64 (2020). https://doi.org/10.1007/s11467-019-0917-0
19. T. Neff, H. Feldmeier, Clustering and other exotic phenomena in nuclei. Eur. Phys. J. Spec. Top **156**, 69–92 (2008). https://doi.org/10.1140/epjst/e2008-00609-y
20. J.A. Wheeler, On the mathematical description of light nuclei by the method of resonating group structure. Phys. Rev. **52**, 1107–1122 (1937). https://doi.org/10.1103/PhysRev.52.1107
21. K. Arai, Y. Suzuki, K. Varga, Neutron-proton halo structure of the 3.563-MeV $0^+$ state in $^{6}$Li. Phys. Rev. C **51**, 2488–2493 (1995). https://doi.org/10.1103/PhysRevC.51.2488
22. K. Arai, K. Katō, S. Aoyama, Di-trinucleon resonance states of $a = 6$ systems in a microscopic cluster model. Phys. Rev. C **74**, 034305 (2006). https://doi.org/10.1103/PhysRevC.74.034305
23. D.L. Hill, J.A. Wheeler, Nuclear constitution and the interpretation of fission phenomena. Phys. Rev. **89**, 1102–1145 (1953). https://doi.org/10.1103/PhysRev.89.1102
24. J.J. Griffin, J.A. Wheeler, Collective motions in nuclei by the method of generator coordinates. Phys. Rev. **108**, 311–327 (1957). https://doi.org/10.1103/PhysRev.108.311
25. N. Hizawa, K. Hagino, K. Yoshida, Generator coordinate method with a conjugate momentum: application to particle number projection. Phys. Rev. C **103**, 034313 (2021). https://doi.org/10.1103/PhysRevC.103.034313
26. C. Jiao, C.W. Johnson, Union of rotational and vibrational modes in generator-coordinate-type calculations, with application to neutrinoless double-$\beta$ decay. Phys. Rev. C **100**, 031303 (2019). https://doi.org/10.1103/PhysRevC.100.031303
27. S. Saito, Interaction between clusters and Pauli principle. Prog. Theor. Phys. **41**, 705–722 (1969). https://doi.org/10.1143/PTP.41.705
28. H. Horiuchi, Y. Kanada-En'yo, Structure of light exotic nuclei studied with amd model. Nucl. Phys. A **616**, 394–405 (1997). https://doi.org/10.1016/S0375-9474(97)00108-5
29. Y. Kanada-En'yo, M. Kimura, H. Horiuchi, Antisymmetrized molecular dynamics: a new insight into the structure of nuclei. Comptes Rendus. Phys. **4**, 497–520 (2003). https://doi.org/10.1016/S1631-0705(03)00062-8
30. Y. Kanada-En'yo, H. Horiuchi, Clustering in yrast states of $^{20}$Ne studied with antisymmetrized molecular dynamics. Prog. Theor. Phys. **93**, 115–136 (1995). https://doi.org/10.1143/ptp/93.1.115
31. A. Tohsaki, H. Horiuchi, P. Schuck et al., Alpha cluster condensation in $^{12}$C and $^{16}$O. Phys. Rev. Lett. **87**, 192501 (2001). https://doi.org/10.1103/PhysRevLett.87.192501
32. P. Ring, P. Schuck, *The nuclear many-body problem* (Springer Science & Business Media, New York, 2004)
33. K. Capelle, Variational calculation of many-body wave functions and energies from density functional theory. J. Chem. Phys. **119**, 1285–1288 (2003). https://doi.org/10.1063/1.1593014
34. O.E. Alon, A.I. Streltsov, L.S. Cederbaum, Interacting fermions and bosons with definite total momentum. Phys. Rev. B **71**, 125113 (2005). https://doi.org/10.1103/PhysRevB.71.125113
35. B. Zhou, Y. Funaki, H. Horiuchi et al., The 5$\alpha$ condensate state in $^{20}$Ne. Nat. Commun. **14**, 8206 (2023). https://doi.org/10.1038/s41467-023-43816-9
36. L. Robledo, T. Rodríguez, R. Rodríguez-Guzmán, Mean field and beyond description of nuclear structure with the gogny force: a review. J. Phys. G Nucl. Part. Phys. **46**, 013001 (2018). https://doi.org/10.1088/1361-6471/aadebd
37. C.W. Wang, B. Zhou, Y.G. Ma, Nonlocalized clustering in $^{18}$O. Eur. Phys. J. A **59**, 49 (2023). https://doi.org/10.1140/epja/s10050-023-00961-5
38. Y.T. Oganessian, V.I. Zagrebaev, J.S. Vaagen, "Di-neutron" configuration of $^{6}$He. Phys. Rev. Lett. **82**, 4996–4999 (1999). https://doi.org/10.1103/PhysRevLett.82.4996
39. B. Danilin, I. Thompson, J. Vaagen et al., Three-body continuum structure and response functions of halo nuclei (i): $^{6}$He. Nucl. Phys. A **632**, 383–416 (1998). https://doi.org/10.1016/S0375-9474(98)00002-5

40. M. Zhukov, B. Danilin, D. Fedorov et al., Bound state properties of borromean halo nuclei: $^{6}$He and $^{11}$Li. Phys. Rep. **231**, 151–199 (1993). https://doi.org/10.1016/0370-1573(93)90141-Y
41. X. Zhang, W. Lin, J.M. Yao et al., Optimization of the generator coordinate method with machine-learning techniques for nuclear spectra and neutrinoless double-$\beta$ decay: Ridge regression for nuclei with axial deformation. Phys. Rev. C **107**, 024304 (2023). https://doi.org/10.1103/PhysRevC.107.024304
42. A.M. Romero, J.M. Yao, B. Bally et al., Application of an efficient generator-coordinate subspace-selection algorithm to neutrinoless double-$\beta$ decay. Phys. Rev. C **104**, 054317 (2021). https://doi.org/10.1103/PhysRevC.104.054317
43. J. Martínez-Larraz, T.R. Rodríguez, Optimization of the number of intrinsic states included in the discrete generator coordinate method. Phys. Rev. C **106**, 054301 (2022). https://doi.org/10.1103/PhysRevC.106.054301
44. K. Varga, Y. Suzuki, R. Lovas, Microscopic multicluster description of neutron-halo nuclei with a stochastic variational method. Nucl. Phys. A **571**, 447–466 (1994). https://doi.org/10.1016/0375-9474(94)90221-6
45. Y. Suzuki, R.G. Lovas, K. Varga, Study of light exotic nuclei with a stochastic variational method: application to lithium isotopes. Prog. Theor. Phys. Suppl. **146**, 413–421 (2002). https://doi.org/10.1143/PTPS.146.413
46. T. Suhara, Y. Kanada-En'yo, Quadrupole deformation $\beta$ and $\gamma$ constraint in a framework of Antisymmetrized molecular dynamics. Prog. Theor. Phys. **123**, 303–325 (2010). https://doi.org/10.1143/PTP.123.303
47. Y. Fukuoka, S. Shinohara, Y. Funaki et al., Deformation and cluster structures in $^{12}$C studied with configuration mixing using skyrme interactions. Phys. Rev. C **88**, 014321 (2013). https://doi.org/10.1103/PhysRevC.88.014321
48. T. Ichikawa, N. Itagaki, Optimization of basis functions for multiconfiguration mixing using the replica exchange Monte Carlo method and its application to $^{12}$C. Phys. Rev. C **105**, 024314 (2022). https://doi.org/10.1103/PhysRevC.105.024314
49. Y.F. Liu, C. Su, J. Liu et al., Improved naive bayesian probability classifier in predictions of nuclear mass. Phys. Rev. C **104**, 014315 (2021). https://doi.org/10.1103/PhysRevC.104.014315
50. X.C. Ming, H.F. Zhang, R.R. Xu et al., Nuclear mass based on the multi-task learning neural network method. Nucl. Sci. Tech. **33**, 48 (2022). https://doi.org/10.1007/s41365-022-01031-z
51. Z.P. Gao, Y.J. Wang, H.L. Lü et al., Machine learning the nuclear mass. Nucl. Sci. Tech. **32**, 109 (2021). https://doi.org/10.1007/s41365-021-00956-1
52. Y.Y. Cao, J.Y. Guo, B. Zhou, Predictions of nuclear charge radii based on the convolutional neural network. Nucl. Sci. Tech. **34**, 152 (2023). https://doi.org/10.1007/s41365-023-01308-x
53. Z.Y. Yuan, D. Bai, Z.Z. Ren et al., Theoretical predictions on $\alpha$-decay properties of some unknown neutron-deficient actinide nuclei using machine learning. Chin. Phys. C **46**, 024101 (2022). https://doi.org/10.1088/1674-1137/ac321c
54. G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks. Science **355**, 602–606 (2017). https://doi.org/10.1126/science.aag2302
55. X. Zhang, W. Lin, J.M. Yao et al., Optimization of the generator coordinate method with machine-learning techniques for nuclear spectra and neutrinoless double-$\beta$ decay: Ridge regression for nuclei with axial deformation. Phys. Rev. C **107**, 024304 (2023). https://doi.org/10.1103/PhysRevC.107.024304
56. J. Keeble, A. Rios, Machine learning the deuteron. Phys. Lett. B **809**, 135743 (2020). https://doi.org/10.1016/j.physletb.2020.135743
57. N.P. Harrington, D. Branford, K. Föhl et al., Experimental study of the halo nucleus $^{6}$He using the $^{6}$Li$(\gamma,\pi^{+})^{6}$He reaction. Phys. Rev. C **75**, 044311 (2007). https://doi.org/10.1103/PhysRevC.75.044311
58. D.P. Kingma, J. Ba, Adam: a method for stochastic optimization. arXiv preprint . https://doi.org/10.48550/arXiv.1412.6980
59. V.S. Vasilevsky, K. Katō, N. Takibayev, Systematic investigation of the hoyle-analog states in light nuclei. Phys. Rev. C **98**, 024325 (2018). https://doi.org/10.1103/PhysRevC.98.024325
60. N. Itagaki, A. Kobayakawa, S. Aoyama, New description of light nuclei by extending the AMD approach. Phys. Rev. C **68**, 054302 (2003). https://doi.org/10.1103/PhysRevC.68.054302
61. T. Furumoto, T. Suhara, N. Itagaki, Effect of channel coupling on the elastic scattering of lithium isotopes. Phys. Rev. C **97**, 044602 (2018). https://doi.org/10.1103/PhysRevC.97.044602
62. A. Volkov, Equilibrium deformation calculations of the ground state energies of 1p shell nuclei. Nucl. Phys. **74**, 33–58 (1965). https://doi.org/10.1016/0029-5582(65)90244-0
63. T. Furumoto, T. Suhara, N. Itagaki, Effect of channel coupling on the elastic scattering of lithium isotopes. Phys. Rev. C **97**, 044602 (2018). https://doi.org/10.1103/PhysRevC.97.044602
64. R. Tamagaki, Potential models of nuclear forces at small distances. Prog. Theor. Phys. **39**, 91–107 (1968). https://doi.org/10.1143/PTP.39.91
65. N. Yamaguchi, T. Kasahara, S. Nagata et al., Effective interaction with three-body effects. Prog. Theor. Phys. **62**, 1018–1034 (1979). https://doi.org/10.1143/PTP.62.1018
66. G. Marsaglia, T.A. Bray, A convenient method for generating normal variables. SIAM Rev. **6**, 260–264 (1964). https://doi.org/10.1137/1006063