Optimization of the generator coordinate method with machine-learning techniques for nuclear spectra and neutrinoless double-β decay: Ridge regression for nuclei with axial deformation

X. Zhang,[1] W. Lin,[1] J. M. Yao [ORCID],[1,*] C. F. Jiao [ORCID],[1] A. M. Romero,[2,3,4] T. R. Rodríguez [ORCID],[5,6,7] and H. Hergert[8,9]

[1]*School of Physics and Astronomy, Sun Yat-sen University, Zhuhai 519082, People's Republic of China*
[2]*Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27516-3255, USA*
[3]*Departament de Física Quàntica i Astrofísica (FQA), Universitat de Barcelona (UB), c. Martí i Franqués, 1, 08028 Barcelona, Spain*
[4]*Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), c. Martí i Franqués, 1, 08028 Barcelona, Spain*
[5]*Departamento de Estructura de la Materia, Física Térmica y Electrónica, Universidad Complutense de Madrid, E-28040 Madrid, Spain*
[6]*Departamento de Física Teórica, Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[7]*Centro de Investigación Avanzada en Física Fundamental-CIAFF-UAM, E-28049 Madrid, Spain*
[8]*Facility for Rare Isotope Beams, Michigan State University, East Lansing, Michigan 48824-1321, USA*
[9]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824-1321, USA*

Background: The generator coordinate method (GCM) is an important tool of choice for modeling large-amplitude collective motion in atomic nuclei. Recently, it has attracted increasing interest as it can be exploited to extend *ab initio* methods to the collective excitations of medium-mass and heavy deformed nuclei, as well as the nuclear matrix elements (NME) of candidates for neutrinoless double-beta ($0\nu\beta\beta$) decay.

Purpose: The computational complexity of the GCM increases rapidly with the number of collective coordinates. It imposes a strong restriction on the applicability of the method. We aim to exploit statistical machine-learning (ML) algorithms to speed up GCM calculations and ultimately provide a more efficient description of nuclear energy spectra and other observables such as the NME of $0\nu\beta\beta$ decay without loss of accuracy.

Method: In this work, we propose a subspace-reduction algorithm that employs optimal statistical ML models as surrogates for exact quantum-number projection calculations for norm and Hamiltonian kernels. The model space of the original GCM is reduced to a subspace relevant for nuclear low energy spectra and the NME of ground state to ground state $0\nu\beta\beta$ decay based on the orthogonality condition (OC) and the energy-transition-orthogonality procedure (ENTROP), respectively. Nuclear energy spectra are determined by the GCM through the configuration mixing within this subspace. For simplicity, the polynomial ridge regression (RR) algorithm is used to learn the norm and Hamiltonian kernels of axially deformed configurations. The efficiency and accuracy of this algorithm are illustrated for ^{76}Ge and ^{76}Se by comparing results obtained using the optimal RR models to direct GCM calculations. The nonrelativistic Gogny force D1S and relativistic energy density functional PC-PK1, a valence-space shell-model Hamiltonian, and a modern nuclear interaction derived from chiral effective field theory are employed.

Results: The low-lying energy spectra of ^{76}Ge and ^{76}Se, as well as the $0\nu\beta\beta$-decay NME between their ground states, are computed. The results show that the performance of the GCM+OC/ENTROP+RR is more robust than that of the GCM + RR alone, and the former can reproduce the results of the original GCM calculation accurately with a significantly reduced computational cost.

Conclusions: Statistical ML algorithms, when implemented properly, can accelerate GCM calculations without loss of accuracy. In applications with axially deformed states, the computation time can be reduced by a factor of three to nine for energy spectra and NMEs, respectively. This factor is expected to increase significantly with the number of employed generator coordinates.

I. INTRODUCTION

The core idea of the generator coordinate method (GCM) is that the wave functions of nuclear states can be represented as a superposition of a set of nonorthogonal basis functions, such as Slater determinants, that are generated by some continuously changing parameters called generator coordinates [1,2]. In practical applications, one often chooses macroscopic quantities that define global nuclear properties such as deformation, and discretizes the associated coordinates on meshes. The dynamics are then described by the Ritz variational principle, where the variation is usually performed with respect to the expansion coefficients in the chosen basis, and

---

*Corresponding author: yaojm8@sysu.edu.cn

sometimes the basis configurations themselves [3,4], resulting in a generalized eigenvalue equation (GEE). In this way, the GCM provides a general approach for solving many-body problems in both nuclear physics [3] and quantum chemistry [5–7] due to the great flexibility of choosing basis functions or generator coordinates. In nuclear physics, the GCM combined with quantum number projections has been extensively employed in studies of the energies and transition rates of low-lying states (see, for instance, Refs. [8–12]). State-of-the-art GCM applications range from the structure of nuclei with triaxial deformation [13–16], to quadrupole-octupole deformed even-even nuclei [17–19], certain odd-mass nuclei [14,20] to the the computation of nuclear matrix elements (NMEs) of $0\nu\beta\beta$ decay [17,21–27]. The latter are vital for interpreting and planning the current- and next-generation tonne-scale experiments for $0\nu\beta\beta$ decays (see the recent reviews [28–30]). GCM calculations most frequently use modern energy density functionals (EDFs) and effective Hamiltonians as inputs, but there have been several works that employ nuclear forces from chiral effective field theory (EFT) in recent years, as the GCM has attracted interest as an pathway for extending nuclear *ab initio* calculations to deformed nuclei [26,27,31–34].

The exact wave functions of nuclear states can in principle be well represented with the GCM ansatz if one chooses a sufficient number of generator coordinates, but this comes at the price of increasing both complexity and computational time. It makes the problem hard to handle exactly because the kernels in the GEE usually require multidimensional integrals of overlap functions over the collective coordinates. For this reason, it is a challenge to extend multidimensional GCM to atomic nuclei throughout the nuclear chart. Practical applications are usually limited to only one or two generator coordinates [35,36]. Therefore, a good choice of generator coordinates or a subset defined by the basis functions becomes important, and this choice is usually based on an *educated guess*, unfortunately.

In applications, one often observes that many of the basis functions connected by the generator coordinates have little contribution to the wave functions of low-lying states and can therefore be safely omitted—see, for instance, Refs. [31,37,38]. In other words, a careful selection of the basis functions can reduce the dimensions of the GEE, and therefore the computational cost. It is worth noting that similar considerations apply to the eigenvector continuation (EC) method [39,40], which finds the eigenvalues and eigenvectors of a Hamiltonian with one or more control parameters, usually the coupling constants of chiral Hamiltonians. In this context, the EC method has been extensively applied to emulate few- and many-body calculations for nuclear structure and scattering [41–45] in recent years. These parameters of EC are analogous to the generator coordinates in GCM, hence finding an efficient way for sampling the basis functions that define a subspace to represent the states of interest is important for both EC and GCM [37]. Several algorithms have been proposed, including the variation-after-projection algorithm [46–48], the stochastic sampling with Monte Carlo techniques [49–51], the choice of low-lying quasiparticle Tamm-Dancoff modes [52], the

energy-transition-orthogonality procedure (ENTROP) [31], and the discrete nonorthogonal shell model (DNO-SM) [53]. We note that many of these algorithms still require a substantial computational effort for the subspace determination.

In the past decade, machine learning (ML) techniques combined with statistical methods have been applied to a variety of nuclear physics problems, ranging from the smallest constituents of matter to the physics of dense astronomical objects—see, for instance, the recent review [54] and the references therein. In some of these applications, a specific statistical ML model is trained to predict nuclear observables directly, including nuclear masses [55–59], charge radii [60–62], β-decay half-lives [63,64], fission yields [65], and many others. In other approaches, statistical ML can be used to enhance nuclear many-body calculations as a surrogate model for expensive computational steps. Examples are the use of a deep neural network committee to optimize collective Hamiltonians for low-lying nuclear states [66], or back-propagation neural networks [67] and kernel Ridge regression [68] to determine density profiles as inputs for nuclear radii and binding energies in the framework of density functional theory.

In the present work, we present the first application of statistical ML techniques to optimize GCM calculations for the low-lying states and the NMEs of $0\nu\beta\beta$ decay of realistic candidate nuclei. The norm and Hamiltonian kernels will be learned by a statistical model. Here, we are confronted with several challenges originating from the norm kernels, which are nonlocal in the collective coordinate space and vary by several orders of magnitude. The solutions of GCM are sensitive to any noise in the norm kernels, which may affect the linear dependence among basis functions and their nontrivial coherence with Hamiltonian kernels, and potentially spoil the GCM description entirely. To tackle these challenges, we propose a subspace reduction algorithm, that uses statistical ML models together with an orthogonality condition (OC) method as an efficient tool to determine the subspace in which the wave functions of nuclear low-lying states can be well represented. Its efficiency and accuracy are illustrated for nuclear energy spectra and $0\nu\beta\beta$ decay NMEs starting from two different EDFs, a valence-space shell-model interaction, and a realistic two- plus three-nucleon Hamiltonian from chiral EFT.

The article is organized as follows. In Sec. II A, the formalism for the GCM, the NME of $0\nu\beta\beta$ decay, and the statistical ML model are introduced. In Sec. III, the performance of the subspace reduction algorithm is illustrated with two different EDFs, and two different Hamiltonians. A summary of our findings and an outlook are given in Sec. IV.

II. FORMALISM

A. The generator coordinate method

In a GCM calculation with quantum-number projections, the nuclear wave function is constructed as follows:

$$\left| \Psi_\alpha^{JMNZ} \right\rangle = \sum_{q=1}^{N_q} \sum_{K=-J}^{J} f_K^{J\alpha}(\boldsymbol{q}) \left| J(MK)NZ, \boldsymbol{q} \right\rangle , \quad (1)$$

where $\alpha = 1, 2, \ldots$ distinguishes the states with the same angular momentum J and the numbers of nucleons. The basis functions $|J(MK)NZ, \boldsymbol{q}\rangle$ are quantum-number projected Hartree-Fock-Bogoliubov (HFB) states labeled by the generator coordinate $\boldsymbol{q}$,

$$|J(MK)NZ, \boldsymbol{q}\rangle = \hat{P}^J_{MK} \hat{P}^N \hat{P}^Z |\Phi(\boldsymbol{q})\rangle. \quad (2)$$

The operator $\hat{P}^J_{MK}$ extracts from the HFB wave function $|\Phi(\boldsymbol{q})\rangle$ the component whose angular momentum along the intrinsic z axis is given by K. The $\hat{P}^{N,Z}$ are the particle number projection (PNP) operators that extract the component with the appropriate neutron number N and proton number Z, respectively. The weight function $f^{J\alpha}_K(\boldsymbol{q})$ in the GCM states given by Eq. (1) is determined by the variational principle, which leads to the discretized Hill-Wheeler-Griffin (HWG) equation [1,2],

$$\sum_{\boldsymbol{q}'K'} \left[\mathcal{H}^J_{KK'}(\boldsymbol{q}, \boldsymbol{q}') - E^J_\alpha \mathcal{N}^J_{KK'}(\boldsymbol{q}, \boldsymbol{q}') \right] f^{J\alpha}_{K'}(\boldsymbol{q}') = 0. \quad (3)$$

The Hamiltonian and norm kernels $\mathcal{H}$ and $\mathcal{N}$ are defined as

$$\mathcal{H}^J_{KK'}(\boldsymbol{q}, \boldsymbol{q}') \equiv \langle J(MK)NZ, \boldsymbol{q}| \hat{\mathbf{H}} |J(MK')NZ, \boldsymbol{q}'\rangle, \quad (4a)$$

$$\mathcal{N}^J_{KK'}(\boldsymbol{q}, \boldsymbol{q}') \equiv \langle J(MK)NZ, \boldsymbol{q}| \hat{\mathbb{1}} |J(MK')NZ, \boldsymbol{q}'\rangle, \quad (4b)$$

where $\hat{\mathbf{H}}$ and $\hat{\mathbb{1}}$ are the Hamiltonian and identity operators, respectively. We note that in the subsequent EDF-based calculations, the Hamiltonian kernels are evaluated based on the *mixed density prescription*: The Hamiltonian overlaps between two different configurations are replaced with the energy which is a functional of mixed densities and currents defined by the two configurations [11].

The HWG equation (3) is solved as follows. We first diagonalize the norm kernel matrix $\mathcal{N}^J_{KK'}(\boldsymbol{q}, \boldsymbol{q}')$ and use its eigenvalues and eigenvectors to construct a set of orthonormal bases $\{|k\rangle\}$, called 'natural states'. To remove the overcompleteness of the original basis functions $|J(MK)NZ, \boldsymbol{q}\rangle$ that stems from the use of continuous quantum numbers, only eigenvectors whose corresponding eigenvalue is larger than a chosen cutoff are included. Then we evaluate the elements of the Hamiltonian matrix in this new subspace $H_{kk'}$, whose eigenvalues then define the energies E^{JNZ}_α of the GCM states. The corresponding eigenvectors are used to determine the weight function $f^{J\alpha}_K(\boldsymbol{q})$. More details can be found in Refs. [3,29], for instance.

B. The nuclear matrix element of $0\nu\beta\beta$ decay

Here, we only consider the NME of $0\nu\beta\beta$ decay corresponding to the transition from the ground state of an even-even nucleus to that of a neighboring even-even nucleus. The spin-parity of both ground states is 0^+. Their wave functions are given by Eq. (1) with $J = K = 0$. For convenience, we simply use the symbol $f(\mathbf{q})$ to replace $f^{J=0,\alpha=1}_{K=0}(\boldsymbol{q})$ and $\mathcal{N}$ for $\mathcal{N}^{J=0}_{K=0,K'=0}$. Therefore, one finds the following expression for the NME:

$$M^{0\nu} = \sum_{\mathbf{q}_F, \mathbf{q}_I} f^*(\mathbf{q}_F) f(\mathbf{q}_I) \langle \Phi(\mathbf{q}_F)| \hat{O}^{0\nu} \hat{P}^{N_I} \hat{P}^{Z_I} \hat{P}^{J=0} |\Phi(\mathbf{q}_I)\rangle = \sum_{\mathbf{q}_F, \mathbf{q}_I} f^*(\mathbf{q}_F) f(\mathbf{q}_I) \mathcal{N}^{1/2}(\boldsymbol{q}_I, \boldsymbol{q}_I) \mathcal{N}^{1/2}(\boldsymbol{q}_F, \boldsymbol{q}_F) \mathcal{M}^{0\nu}(\boldsymbol{q}_F, \boldsymbol{q}_I), \quad (5)$$

where the kernel $\mathcal{M}^{0\nu}(\boldsymbol{q}_F, \boldsymbol{q}_I)$ for the NME can be computed with the help of two-body transition matrix elements $O^{0\nu}_{pp'nn'}$ in a spherical harmonic oscillator basis,

$$\mathcal{M}^{0\nu}(\boldsymbol{q}_F, \boldsymbol{q}_I) = \frac{1}{4} \sum_{pp'nn'} O^{0\nu}_{pp'nn'} \rho_{pp'nn'}(\boldsymbol{q}_F, \boldsymbol{q}_I). \quad (6)$$

The two-body transition density $\rho_{pp'nn'}(\boldsymbol{q}_F, \boldsymbol{q}_I)$ is determined by

$$\rho_{pp'nn'}(\boldsymbol{q}_F, \boldsymbol{q}_I) = \frac{1}{\mathcal{N}^{1/2}(\boldsymbol{q}_I, \boldsymbol{q}_I) \mathcal{N}^{1/2}(\boldsymbol{q}_F, \boldsymbol{q}_F)} \int_0^{2\pi} \frac{e^{-iN_I\varphi_N}}{2\pi} d\varphi_N \int_0^{2\pi} \frac{e^{-iZ_I\varphi_Z}}{2\pi} d\varphi_Z \int d\Omega \, \langle \Phi(\boldsymbol{q}_F)| c^\dagger_p c^\dagger_{p'} c_{n'} c_n e^{i\varphi_N \hat{N}} e^{i\varphi_Z \hat{Z}} \hat{R}(\Omega) |\Phi(\boldsymbol{q}_I)\rangle,$$

$$(7)$$

where the rotation operator reads $\hat{R}(\Omega) = e^{i\varphi\hat{J}_z} e^{i\theta\hat{J}_y} e^{i\psi\hat{J}_z}$, (φ, θ, ψ) are the three Euler angles, and $\varphi_{N/Z}$ are the gauge angles that define the projection operators $\hat{P}^{N/Z}$. The symbols N_I, Z_I are the neutron and proton numbers of the initial nucleus. The $|\Phi(\boldsymbol{q}_{I/F})\rangle$ are the HFB wave functions for the initial and final nucleus, respectively. In the present work, only the long-range transition operator $O^{0\nu}$ in the standard mechanism of light Majorana neutrino exchange is considered. See Ref. [29] for details.

C. Learning kernels with the polynomial ridge regression

For the present proof-of-concept study, we only consider the quadrupole deformation parameter β as a generator co-

ordinate $\mathbf{q}$ in Eq. (1), and we will switch labels in the basis configurations accordingly in the following discussion. Because of this specialization, we have $K = 0$ and the norm kernel simplifies into the following form:

$$\mathcal{N}^J_{00}(\beta, \beta') = \langle J(MK = 0)NZ, \beta| \hat{\mathbb{1}} |J(MK' = 0)NZ, \beta'\rangle$$

$$= \frac{2J + 1}{2} \int_0^\pi d\theta d^J_{00}(\theta) \langle \Phi(\beta)|$$

$$\times e^{i\theta\hat{J}_y} \hat{P}^Z \hat{P}^N |\Phi(\beta')\rangle, \quad (8)$$

where $d^J_{00}(\theta) = \langle J0|e^{i\theta\hat{J}_y}|J0\rangle$ is the Wigner (small) d matrix. This norm kernel will be learned with a statistical ML model in the following.

Previous GCM studies have established the following features of the norm kernel:

(i) For the diagonal element of the norm kernel with $\beta = \beta'$ but without PNP operators, the overlap function can be well approximated with a Gaussian function [3,69,70],

$$\langle \Phi(\beta)| \, e^{i\theta \hat{J}_y} \, |\Phi(\beta)\rangle \simeq \exp\left(-\frac{\theta^2}{2} \, \langle \Phi(\beta)| \, \hat{J}_y^2 \, |\Phi(\beta)\rangle\right). \tag{9}$$

The nondiagonal overlap with $\beta \neq \beta'$ can be parametrized with an extension of the Gaussian overlap approximation [71,72].

(ii) For the nondiagonal element with $\beta \neq \beta'$ and without any projections, the norm kernel can also be approximated with a Gaussian function [3,4,73]

$$\mathcal{N}(\beta, \beta') \simeq \exp\left[-\frac{\gamma(q)}{2}(\beta - \beta')^2\right], \tag{10}$$

where $\gamma(q)$ is a function of the deformation parameter $q = (\beta + \beta')/2$ that can be calculated using the corresponding HFB wave functions.

Of course, the actual norm kernels with the projections of particle-number and angular momentum are expected to have a much more complicated expression that is to be learned by statistical ML models.

In light of the Gaussian structure that was found in the aforementioned results, it is reasonable to attempt to train the logarithm of the norm kernels to avoid dealing with data that spans several orders of magnitude. We expect this logarithm to be well approximated with a polynomial function of the quadrupole deformation parameters β and β'. Therefore, a basic machine-learning algorithm, i.e., *the polynomial ridge regression (RR)* [74,75] is adopted for our purpose. As a test, we have validated that this algorithm exactly reproduces the kernels of a simple model with quadratic approximation [2]. We have also employed other algorithms such as support vector regression, which yielded similar results, but at the cost of significantly increased training time.

Here, we present some details on our implementation of the polynomial RR model. In the data preparation, we compute all the kernels $\mathcal{N}^J(\beta, \beta')$ and $\mathcal{H}^J(\beta, \beta')$ exactly with Eq. (4) for a specific nucleus (Z, A) and a given interval in β large enough to include all the relevant configurations. There are N_q^2 norm kernels and N_q^2 Hamiltonian kernels, where N_q is the number of mesh points in the deformation parameter β. For simplicity, these mesh points are equally distributed with step size $\Delta\beta$. Among all the kernels, those equally distributed in the entire deformation space with the step size of $2\Delta\beta$ or $4\Delta\beta$ are selected as a training set, while the remaining data are used for testing and validation. The number of training data and the number of test data are denoted M_{train} and M_{test}, respectively. Thus, we have the relation $N_q^2 = M_{\text{train}} + M_{\text{test}}$.

The ith predicted value $\hat{y}^{(i)}$ in the polynomial RR is given by the following hypothetical N-degree polynomial function:

$$\hat{y}^{(i)}(\boldsymbol{\theta}; N) = \boldsymbol{X}_N^{(i)} \boldsymbol{\theta}_N, \tag{11}$$

where the ith input vector $\boldsymbol{X}_N^{(i)}$ is defined as

$$\boldsymbol{X}_N^{(i)} = (1, \beta, \beta', \cdots, \beta^N, \beta^{N-1}\beta', \cdots, \beta\beta'^{N-1}, \beta'^N)^{(i)} \tag{12}$$

with $(N+2)(N+1)/2$ features, and the corresponding weight parameters

$$\boldsymbol{\theta}_N^T = \left(\theta_0^0, \theta_0^1, \theta_1^0, \cdots, \theta_0^N, \theta_1^{N-1} \cdots \theta_{N-1}^1, \theta_N\right). \tag{13}$$

The degree N of the polynomial is a hyperparameter controlling the complexity of the model. We note from Eq. (12) that the polynomial RR is more flexible than the ridge regression with the N-degree polynomial kernel [74,75]. Besides, one can see that the polynomial RR fits the single, high-degree polynomial function (11) to all of the values in the training set. Compared to the spline interpolation which fits low-degree polynomials to small subsets of the training set, the polynomial RR works equally well when the degree of polynomials is chosen appropriately. However, the polynomial RR is easier to be extended to the case with multiple coordinates and the possible overfitting problem is avoided by adding the Tikhonov regularization term [74] to the mean-square error (MSE) in the definition of loss function,

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}; N, \alpha) &= \sum_{i=1}^{M_{\text{train}}} [\hat{y}^{(i)}(\boldsymbol{\theta}; N) - y^{(i)}]^2 + \alpha \boldsymbol{\theta}_N^T \boldsymbol{\theta}_N \\
&= (\hat{\boldsymbol{Y}} - \boldsymbol{Y})^T(\hat{\boldsymbol{Y}} - \boldsymbol{Y}) + \alpha \boldsymbol{\theta}_N^T \boldsymbol{\theta}_N,
\end{aligned} \tag{14}$$

where the vector $\boldsymbol{Y}$ is defined by

$$\boldsymbol{Y} = (y^{(1)}, \cdots, y^{(M_{\text{train}})})^T \tag{15}$$

with $y^{(i)}$ chosen as the logarithmic value of the ith norm kernel, and $\hat{y}^{(i)}$ is the output (11) of the RR model. The summation in Eq. (14) runs through all the kernels in the training set.

The minimization of the above loss function leads to a normal equation with the solution given by [75]

$$\boldsymbol{\theta}_N = (X^T X + \alpha \mathbf{I})^{-1} X^T Y, \tag{16}$$

where $\mathbf{I}$ is the identity matrix. The ridge parameter $\alpha \geqslant 0$ controls how much one wants to regularize the model. When the α is zero, the RR is simplified into the linear regression [76]. With the increase of α, the regularization term dominates the squared loss function and the weight coefficients $\boldsymbol{\theta}_N$ tend to be small, providing a way to prevent the overfitting problem. In practice it is necessary to tune α in such a way that a balance is maintained between both. Instead of determining the weight coefficients $\boldsymbol{\theta}_N$ analytically according to Eq. (16) by computing the inverse of the normal matrix $(X^T X + \alpha \mathbf{I})$, the vector parameters $\boldsymbol{\theta}_N$ can also be determined by minimizing the loss function $\mathcal{L}$ with the gradient descent (GD) method for given hyperparameters (N, α). We note that in general both methods give the same solution to the model parameters. Compared to the GD method, the normal-equation method is usually employed in the regression with a small number of features. However, with the increase of the hyperparameter N and the size of training set, the inverse of the normal matrix becomes difficult to compute. In contrast, the computation complexity of the GD method grows moderately with the number of model parameters. In this work, we employ the GD
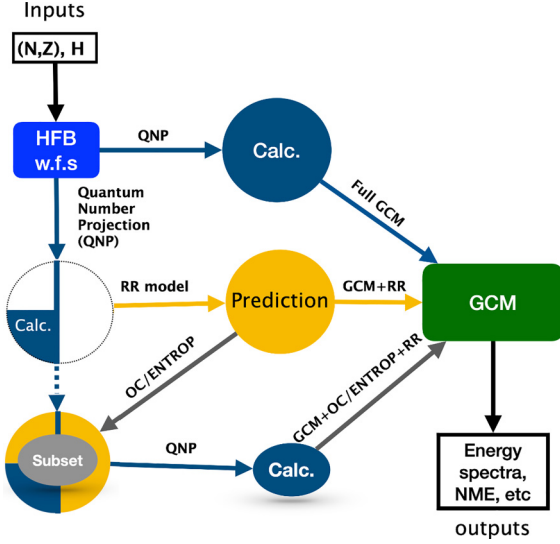
FIG. 1. Comparison of the flowcharts for the full GCM, GCM + RR, and GCM+OC/ENTROP+RR approaches. The area of different shapes represents the domain of kernels obtained from different methods indicated with colored arrows. The vertical line indicates the diagonal elements of kernels. See text for details.

method to determine the weight parameters in the polynomial RR. More detailed introduction to ML models in physics can be found in, for instance, Ref. [77].

III. ILLUSTRATIONS

In the following, we will study three different procedures for implementing the polynomial RR into GCM calculations, starting from two EDFs, a shell-model interaction, and a chiral $NN + 3N$ interaction. They are illustrated with a flow chart shown in Fig. 1. In the GCM + RR, a small portion (about 1/4) of the norm and Hamiltonian kernels are calculated exactly with quantum number projection (QNP) method, providing inputs for training the RR model parameters. Once the model parameters are optimized, they are used to predict all the norm and Hamiltonian kernels which serve as inputs of the GCM calculation. In the GCM + RR procedure, the noise introduced by the RR model is expected to generate errors in the final results. To mitigate these errors, we propose a combination of GCM, OC/ENTROP, and RR, in which a subset is selected based on either the OC method or ENTROP algorithm. An exact GCM calculation is carried out within this subset afterward.

The procedure of the GCM + OC + RR method for nuclear low-lying states is as follows:

(i) All the configurations $|\Phi(\beta)\rangle$ are sorted by their projected energies, i.e., the ratios of the kernels $\mathcal{H}^J_{00}(\beta, \beta)/\mathcal{N}^J_{00}(\beta, \beta)$ for the nucleus of interest. To this end, one needs to calculate these ratios for all the configurations, which requires $O(N_q)$ computational effort. We note that the ordering of the configurations is different for different angular momenta J.

(ii) Starting from the configuration with the lowest energy and stepping from low to high energy, the $(n + 1)$th

configuration $|n + 1\rangle$ is added into the subspace if its projection onto the subspace spanned by the already selected n configurations, defined by

$$
\begin{aligned}
L(n, n + 1) &= \frac{\langle n + 1|P^{(n)}|n + 1\rangle}{\langle n + 1 \mid n + 1\rangle} \\
&= \frac{\boldsymbol{\gamma}^{(n)\dagger}(\boldsymbol{S}^{(n)})^{-1}\boldsymbol{\gamma}^{(n)}}{\langle n + 1 \mid n + 1 \rangle},
\end{aligned} \tag{17}
$$

is smaller than a pre-selected cutoff parameter L_c [31]. This implies that the new configuration is approximately orthogonal to the previous configurations, hence the name of this stage. In Eq. (17), $S^{(n)}_{ij} = \langle i \mid j \rangle$ and $\gamma^{(n)}_i = \langle i \mid n + 1 \rangle$ are nothing but the matrix elements of the norm kernel $\mathcal{N}^J_{00}(\beta, \beta')$. Using the orthogonality criterion, a subspace $\mathcal{S}_{L_c} = \{|1\rangle, |2\rangle, \ldots, |n\rangle\}_{L_c}$ is eventually determined for a given value of L_c. In practice, we employ the pre-calculated norm kernels by the QNP method with the rest given by the RR model in the subspace-selection procedure.

(iii) The norm kernels and Hamiltonian kernels for the configurations within the subspace $\mathcal{S}_{L_c}$ are determined by the QNP method and they are used in the solution of the HWG equation (3). The convergence of each observable against the cutoff parameter L_c is examined.

In the calculation of the NME of $0\nu\beta\beta$ decay, for comparison, we also employ the recently developed ENTROP algorithm to select the subspaces for the initial and the final nuclei simultaneously. Details about the ENTROP algorithm can be found in Ref. [31].

A. EDF-based GCM calculations

1. A nonrelativistic EDF: Gogny D1S

Figures 2 and 3 show the distributions of the norm kernels $\mathcal{N}^J(\beta, \beta')$ and the ratio of kernels $\mathcal{H}^J(\beta, \beta')/\mathcal{N}^J(\beta, \beta')$ with different angular momentum J from quantum-number-projection calculations for the HFB states based on the Gogny D1S force [78,79]. The norm kernels with different angular momenta J are distributed differently, but they share a common feature in that they are mainly concentrated along the diagonal line with $\beta = \beta'$. For the $J = 0$ case, the norm kernel is dominated by a product of two Gaussian functions centered at $\beta = \beta' = 0$. It can be understood that the spherical state only contains a $J = 0$ component. The distributions of the diagonal element $\mathcal{N}^J_{00}(\beta, \beta)$ with $J \neq 0$ share the similar feature that the peak locates at a deformed state, approximately symmetric with respect to $\beta = 0$. This is a general feature of norm kernels [80]. One may expect that this feature can be well captured by the polynomial regression on their logarithmic values. Nevertheless, as shown in Fig. 3(a) and 3(c), the matrix elements of norm kernels vary by several orders of magnitude in the entire deformation space, which is a challenge for ML algorithms. A small error in the norm kernel may degrade the description of GCM, as discussed in detail later on. In contrast, the variation of the
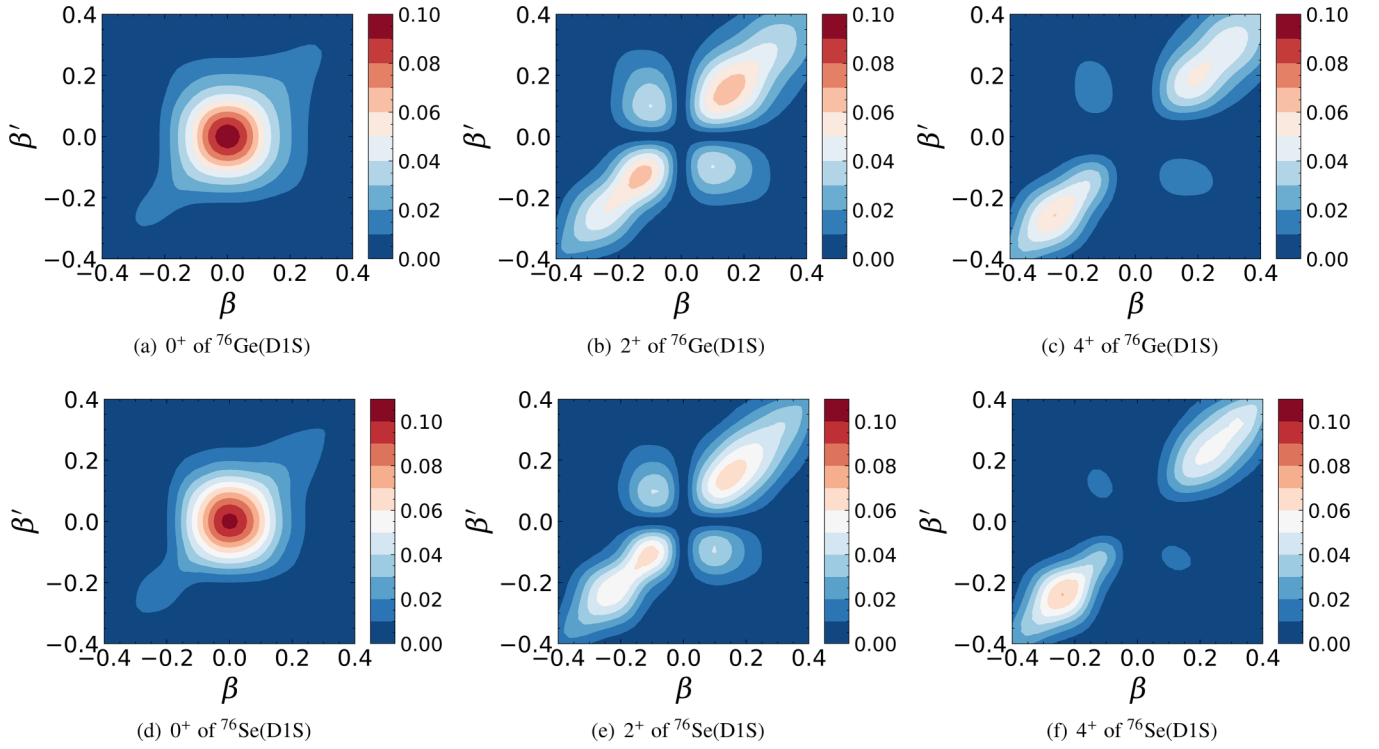
FIG. 2. The distribution of norm kernels $\mathcal{N}_{00}^{J}(\beta, \beta')$ for ^{76}Ge (a), (b), (c) and ^{76}Se (d), (e), (f) from the HFB calculation using the Gogny D1S force as a function of the quadrupole deformation parameters β, β', where the angular momentum is $J = 0$ (a), (d), 2 (b), (e), and 4 (c), (f), respectively.

ratios $\mathcal{H}_{00}^{J=0}(\beta, \beta')/\mathcal{N}_{00}^{J=0}(\beta, \beta')$ with β and β' is moderate and thus expected to be more readily learned by ML algorithms.

The parameters of our polynomial RR models for norm kernels $\mathcal{N}$ and the ratios $\mathcal{H}/\mathcal{N}$ are optimized as explained

FIG. 3. The distributions of norm kernels $\ln(\mathcal{N})$ and the ratio of kernels $\mathcal{H}(\beta, \beta')/\mathcal{N}(\beta, \beta')$ with $J = 0$ for ^{76}Ge (a), (b) and ^{76}Se (c), (d) as a function of the quadrupole deformation parameters (β, β') from the HFB calculation using the Gogny D1S force.

in Sec. II C. Our findings in the training processes are as follows:

(i) The optimal RR model captures the Hermiticity of the kernels, even though it is not strictly enforced in the model's construction at present.

(ii) Selecting training data that are uniformly distributed in the (β, β') plane usually leads to a smaller MSE in Eq. (14) than random sampling. We note that this approach does not scale well to larger numbers of generator coordinates, for which a more robust sampling strategy needs to be explored in the future.

(iii) As shown in Figs. 2 and 3, the norm kernels with $J = 0$ of strongly deformed configurations ($|\beta| > 0.4$) are usually small (less than 10^{-3}). Including these kernels in the training procedure may spoil the description of the model. For kernels with $J \neq 0$, we exclude configurations around the spherical shape with $|\beta| < 0.06$ to guarantee good performance of the model because they are negligible.

Figure 4 shows the RMSE of the kernels $\ln \mathcal{N}$ and the ratios $\mathcal{H}/\mathcal{N}$ with $J = 0$ for both the training set and test set as a function of the degree N of the polynomials in the RR model. For comparison, the results by the RR models with a small (10^{-14}) and large (10^{-1}) value of the ridge parameter α are presented. First, we find that the RR model with $\alpha = 10^{-14}$ works much better than the RR model with $\alpha = 0$, which is the simple linear regression. With the nonzero regularization term in the RR model, the overfitting problem becomes
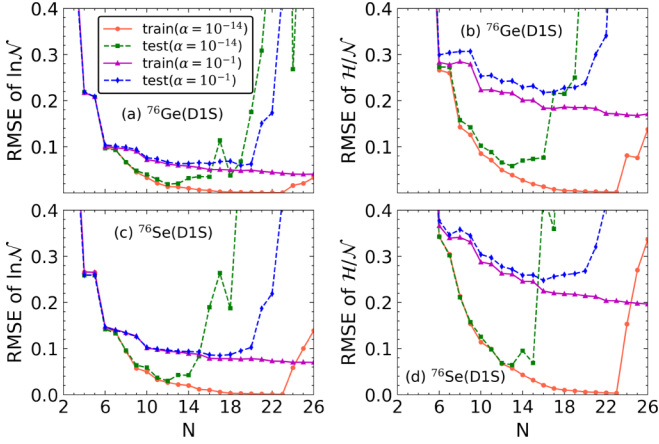
FIG. 4. The root-mean-square error (RMSE) of the RR models for the training set and test set of the norm kernels $\ln \mathcal{N}$ (a), (c) and the ratio of kernels $\mathcal{H}(\beta, \beta')/\mathcal{N}(\beta, \beta')$ (b), (d) with $J = 0$ for ^{76}Ge (a), (b) and ^{76}Se (c), (d) as a function of the degree parameter N of the polynomials, where the results with the ridge parameter α chosen as 10^{-14} and 10^{-1}, respectively, are given for comparison.

moderate, even though it still appears for $N > 12$. Second, with the choice of a larger value of $\alpha(=10^{-1})$, the occurrence of overfitting problem is extended to a larger value of $N(=20)$. Compared to the RR model with $\alpha = 10^{-14}$, the RMSE in the model with $\alpha(=10^{-1})$ is systematically larger. Therefore, in this work, the hyperparameters $N = 12$, $\alpha(=10^{-14})$ are employed in the RR model if not mentioned explicitly. Selecting the value of N giving the best description of the test set, we show the covariance matrix of the RR model for the norm kernels of ^{76}Ge with $J = 0$ in Fig. 5. One can see that the
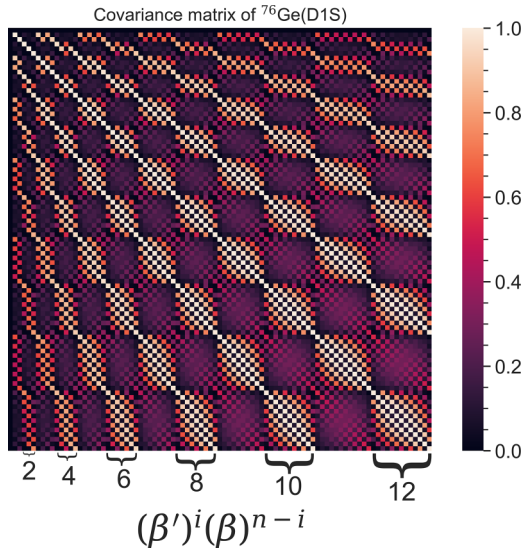
FIG. 5. The covariance matrix of the RR model (14) for the norm kernels of ^{76}Ge with $J = 0$ by the Gogny D1S force, where the degree parameter $N = 12$ and ridge parameter $\alpha = 10^{-14}$. The number of features $(\beta')^i(\beta)^{n-i}$ in Eq. (12) is $(N+2)(N+1)/2$, where the integer number $n \in [0, N]$ is shown in the bottom of the figure with the integer number i varying from 0 to n.

features $(\beta')^i(\beta)^{n-i}$ with the even (odd) values of i and $n(\leqslant N)$ exhibit strong correlation with each other. It indicates that the optimal RR model is still reducible. We will examine the impact of different choices of N on nuclear energy spectra later.

The relative deviations of the norm and Hamiltonian kernels by the optimal RR model are displayed in Fig. 6. The deviation for the norm kernels can be up to 10% for ^{76}Se, while that for the ratio $\mathcal{H}^0/\mathcal{N}^0$ is less than 0.05%. The noise introduced by the RR model into the norm kernels may spoil the correlation relations among different kernels and thus the orthogonality property of different configurations. It is shown in Fig. 7 that the energy plateau (in particular for the excited states of ^{76}Se) becomes slightly worse when the RR-model-predicted kernels, instead of the exactly calculated ones, are used in the GCM calculations. The impact of noise in the kernels of a generalized eigenvalue problem has also been discussed recently in the EC method [81], where a trimmed

FIG. 6. The relative deviation δ of the kernels predicted by the optimal RR model for (a), (b) ^{76}Ge and (c), (d) ^{76}Se, where the δ is defined as $\delta^{(i)} = (\hat{y}^{(i)} - y^{(i)})/y^{(i)}$ with y being the (a), (c) norm kernel $\mathcal{N}$ or (b), (d) the ratio of kernels $\mathcal{H}/\mathcal{N}$, respectively. See main text for details.
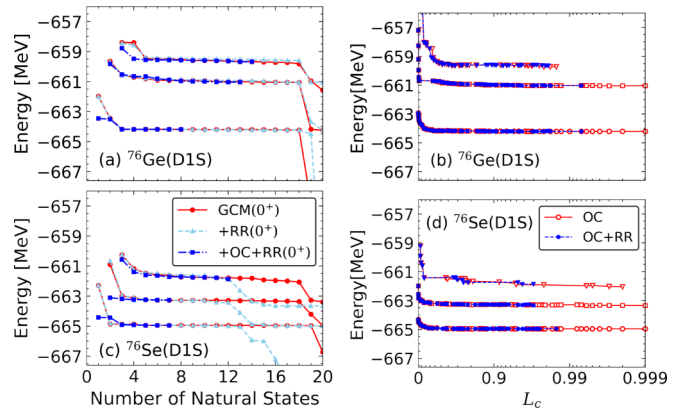

FIG. 7. The energies of states with angular momentum $J = 0$ as a function of (a), (c) the number of natural states and (b), (d) the cutoff parameter L_c from different calculations for (a), (b) ^{76}Ge and (c), (d) ^{76}Se, respectively. See main text for details.

sampling algorithm was proposed to mitigate this issue. We also find that the energy plateau is much worse for the RR model with $\alpha = 0$ (not shown here), demonstrating the important role of the regularization term played in the optimization of GCM with the polynomial RR model.

To mitigate the impact of the noise introduced by the RR model on energy spectra, we employ the OC method to select the subspace $\mathcal{S}_{L_c}$ based on the RR-model-predicted kernels, as discussed before. Figure 7 shows the convergence of the energies of the first three 0^+ states as a function of the cutoff parameter L_c. For comparison, the results from the GCM+OC calculations based on the exactly calculated norm kernels are also given. One can see that the GCM + OC + RR can reasonably reproduce the convergence behavior of the GCM+OC. The value of L_c is determined based on the convergence behavior, which in principle varies with each state. We note that if the same L_c value is taken for all the states, as in the examples discussed later, the performance of OC is slightly worse for the excited states. This can probably be attributed to the fact that the selection of candidate configurations for the subspace follows their energy ordering. The selected subspace is thus expected to be more complete for the ground state than for the excited states. Once the subspace is defined, we calculate both norm kernels and Hamiltonian kernels within this subspace and use them to carry out GCM calculations. Interestingly, but perhaps not unexpectedly, we find that the subspaces for different low-lying states differ from each other only by a few configurations. Fig. 7 shows that the energy of each state terminates at the number of natural states defined by the dimension of the corresponding subspace.

Figure 8 shows the distributions of the eigenvalues of the norm kernels with $J = 0$ from the exact quantum-number projection calculation and from the RR model prediction. The eigenvalues of the norm kernels within the selected subspace for the ground state are also given for comparison. We see that eigenvalues with values smaller than 10^{-3} from the RR model prediction are different from the exactly calculated ones. This explains the degradation of the energy plateaus in Fig. 7. In contrast, the eigenvalues obtained with GCM + OC + RR match the full GCM eigenvalues until the limit of the subspace is reached. Thus, the introduction of noise by the RR model is avoided in the GCM + OC + RR approach.

Figure 9 displays the measure of the distance in the ground-state wave function $||\Psi^{(n)} - \Psi_{\text{GCM}}^{(\text{full})}||$ and in energy, $|E^{(n)} - E_{\text{GCM}}^{(\text{full})}|$ with respect to that by the full GCM calculation (with N_q HFB states) as a function of the number of HFB states in both GCM and GCM + OC + RR calculations for ^{76}Ge and ^{76}Se, where the HFB states are ordered in energy, and the measure of the distance in the wave function is defined as [37]

$$||\Psi^{(n)} - \Psi_{\text{GCM}}^{(\text{full})}|| = \sqrt{2\left[1 - \text{Re}\left(\left\langle \Psi^{(n)} \middle| \Psi_{\text{GCM}}^{(\text{full})} \right\rangle\right)\right]} \qquad (18)$$

with

$$\left\langle \Psi^{(n)} \middle| \Psi_{\text{GCM}}^{(\text{full})} \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{N_q} f^{(n)*}(\beta_i) \mathcal{N}^{J=0}(\beta_i, \beta_j) f_{\text{GCM}}^{(\text{full})}(\beta_j). \qquad (19)$$
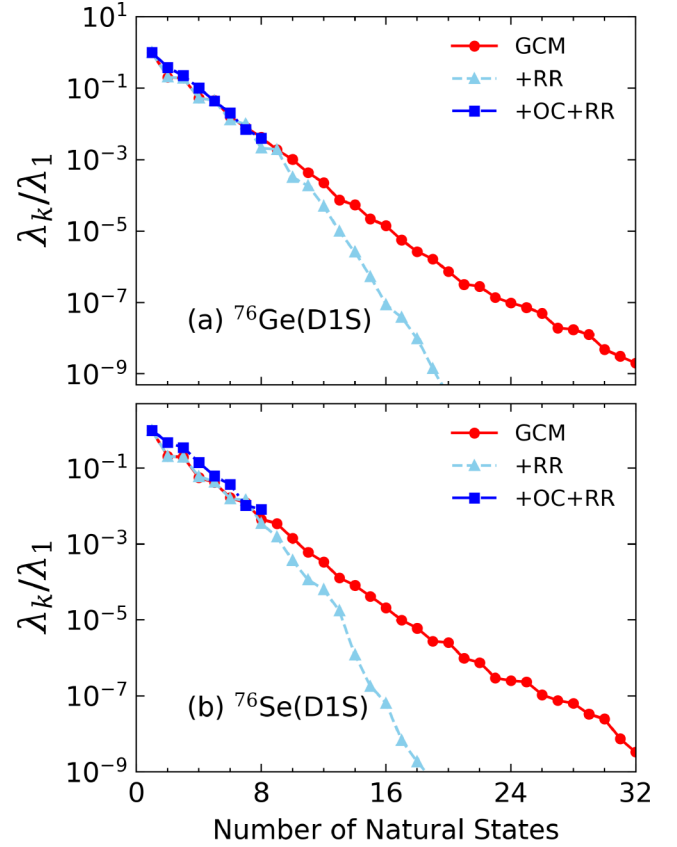

FIG. 8. The distribution of eigenvalues λ_k (normalized to the maximal one) of the norm kernels $\mathcal{N}_{00}^{J=0}(\beta, \beta')$ in the full GCM, GCM + RR, and GCM + OC + RR calculations for ^{76}Ge (a) and ^{76}Se (b), where different subspaces are used in the three calculations.
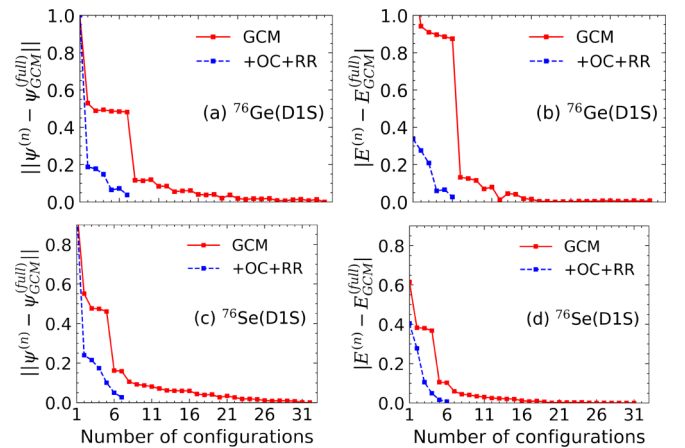

FIG. 9. Comparison of the convergence of the ground-state wave function (a), (c) and energy (b), (d) with respect to that by the full GCM calculation as a function of the number of HFB configurations in both GCM and GCM + OC + RR calculations for (a), (b) ^{76}Ge and (c), (d) ^{76}Se. In the GCM + OC + RR calculation, the L_c value is chosen as 0.872 and 0.824 for (a), (b) ^{76}Ge and (c), (d) ^{76}Se, respectively.
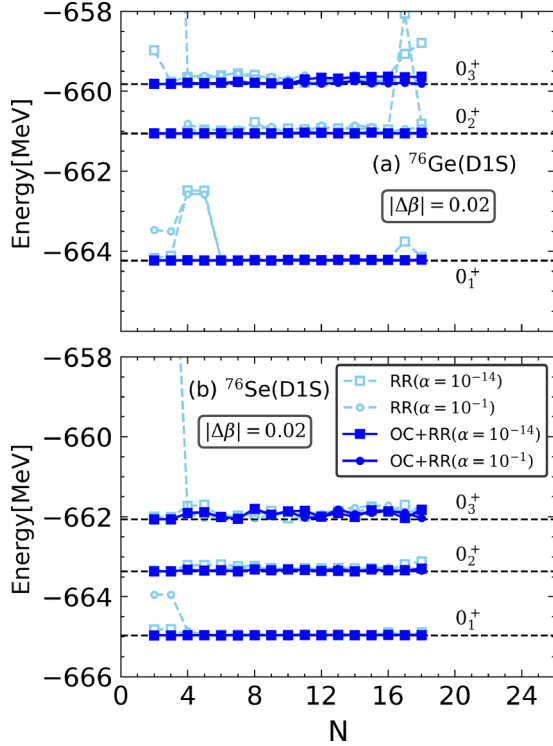
FIG. 10. The energies of the first three 0^+ states for ^{76}Ge (a) and ^{76}Se (b) as a function of the degree parameter N of the polynomials from the GCM + RR and GCM + OC + RR calculations using the D1S force. The black dashed lines indicate the energies from the original GCM calculations using the exactly calculated kernels.

Here, $f^{(n)}(\beta_i)$ and $f_{\mathrm{GCM}}^{(\mathrm{full})}(\beta_j)$ are the weight functions (1) of the ground states in the GCM calculations based on the first n and all N_q HFB states, respectively. One can see that the ground-state wave function converges faster to the wave function of the full GCM calculation in the GCM + OC + RR than that in the pure GCM. In the GCM + OC + RR calculation, the residual norm difference between the subspace-projected wave function and the full GCM solution is typically smaller than 0.05, which corresponds to $\langle \Psi^{(n)} | \Psi_{\mathrm{GCM}}^{(\mathrm{full})} \rangle \simeq 0.999$.

To check how the hyperparameters (N, α) in our RR models affect nuclear energy spectra, we show the energies of the first three 0^+ states from both the GCM + RR and GCM + OC + RR calculations for both ^{76}Ge and ^{76}Se as a function of N in Fig. 10, where two different values of α are employed for comparison. It is shown that the energies are generally stable under the variations of N for all cases, and consistent with the full GCM results. The observed fluctuations occasionally in the results of GCM + RR calculations can be removed when the OC method is implemented additionally. In other words, the GCM + OC + RR method works well for a large range of values for the hyperparameter (N, α).

The low-energy spectra from different calculations are shown in Fig. 11. One can see that the decrease of the step size $\Delta\beta$ in the deformation parameter from 0.08 (total of nine configurations) to 0.02 (total of 33 configurations) only weakly affects the spectrum. Quantitatively, the energy difference in the ground state of ^{76}Ge introduced by the OC + RR is less
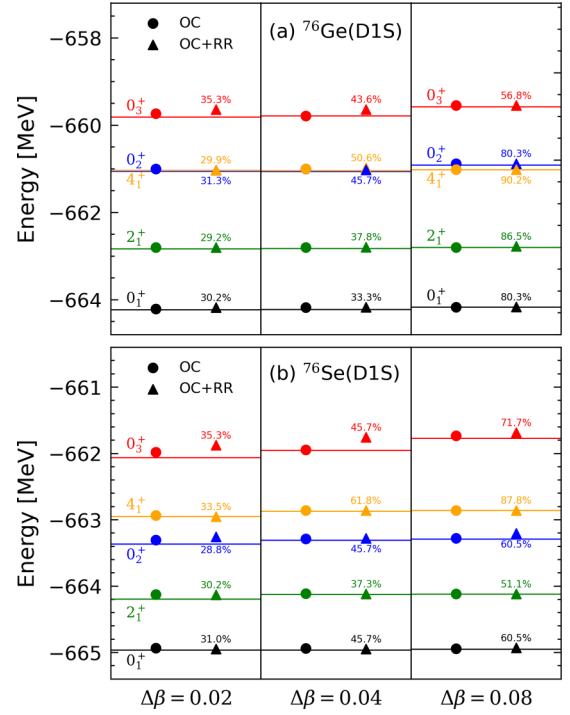

FIG. 11. The energies of low-lying states in ^{76}Ge (a) and ^{76}Se (b) from the GCM (horizontal lines), GCM+OC (circles), and GCM + OC + RR (triangles) calculations, respectively. From left to right columns shows the results from the calculations with the step size $\Delta\beta = 0.02$, $\Delta\beta = 0.04$, and $\Delta\beta = 0.08$, respectively. The number shown nearby each triangle is determined by the ratio $N_S/N_{\mathcal{F}}$, where N_S is the number of kernels for the configurations within the selected subspace for a give L_c and those for the training set, while $N_{\mathcal{F}}$ is the number of kernels for the configurations in the full space.

than 50 keV for all cases. This error is slightly larger for the 0_3^+ state, but it is still around 150 keV for ^{76}Ge and around 200 keV for ^{76}Se. In both nuclei, the energy difference between the GCM + OC + RR and GCM+OC is about 20 keV. In other words, with the application of the OC to GCM calculation, the error introduced by the RR model is negligible. In the current application, the use of the RR model reduces the computational time by a factor of up to three; detailed numbers for each state are included in Fig. 11. As expected, the denser the mesh for discretizing the quadrupole deformation parameter β in the original set, the more computational time one can save with the statistical ML technique.

Next, we examine the performance of the method for the NME of $0\nu\beta\beta$ decay. In addition to the use of the OC method, the ENTROP algorithm [31] is also employed for comparison. As the calculation of the NME of $0\nu\beta\beta$ decay requires the ground-state wave functions of two nuclei, it imposes a more stringent test on the RR model than energy spectra. Figure 12 shows the NMEs of $0\nu\beta\beta$ decay for ^{76}Ge from GCM+ENTROP, GCM + OC + RR, and GCM+ENTROP+RR calculations. The cost of computational time compared to the full GCM calculation is indicated in each case. One can see that the values by the GCM +
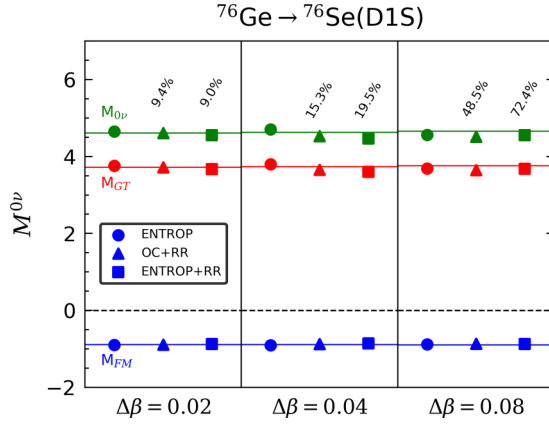
FIG. 12. The Fermi, GT, and total NMEs of $0\nu\beta\beta$ decay for ^{76}Ge. The solid line is obtained from the exact GCM calculation of ^{76}Ge and ^{76}Se, while the circles, triangles, and squares are by GCM+ENTROP, GCM + OC + RR, and GCM+ENTROP+RR calculations, respectively. The numbers represent the percentage of cost in computation time compared to the full GCM calculation using all the configurations, and they are obtained from the multiplication of the ratios for ^{76}Ge and ^{76}Se in Fig. 11. See main text for details.

OC + RR and GCM+ENTROP+RR are generally close to each other, both are slightly different from those by the GCM+ENTROP and full GCM calculations. Quantitatively, the total NME $M^{0\nu}$ from the GCM + OC + RR calculation with $\Delta\beta = 0.02, 0.04, 0.08$ is 4.61, 4.53, and 4.52, respectively, which should be compared to the values 4.61, 4.62, 4.66 from the full GCM calculation without the use of the statistical ML technique. If the subspaces for ^{76}Ge and ^{76}Se are selected based on the ENTROP+RR, the total NME becomes 4.55, 4.46, and 4.55, respectively. The difference in the three numbers is negligibly small. It implies that the NME in the GCM+ENTROP+RR calculation is less sensitive to the choice of the step size in the deformation parameter β than the GCM + OC + RR calculations. In many cases, the predicted NMEs are essentially close to each other and consistent with the value of 4.60 given in Ref. [21].

2. The relativistic EDF: PC-PK1

In this subsection, we present results for ^{76}Ge and ^{76}Se starting from the relativistic EDF PC-PK1 [83]. Again, only axially deformed configurations are employed. The short-range correlation that has been taken into account in Ref. [84] is not included here. More details about the calculations can be found in Ref. [17].

Figure 13 shows the convergence behavior of the energies of states with $J = 0$ in three different calculations with the PC-PK1. Similar to the case of the Gogny D1S, the energy plateaus in the GCM + RR are slightly worse than those by the full GCM with the increase of the number of natural states due to the errors in the kernels introduced by the RR models. With the OC method, the energy plateaus terminate at the number defined by the selected subspace. Figure 14 displays the low-energy spectra for ^{76}Ge from the GCM+OC and GCM + OC + RR calculations. In the $\Delta\beta = 0.2$ case, the energies of states are underestimated evidently by the GCM +
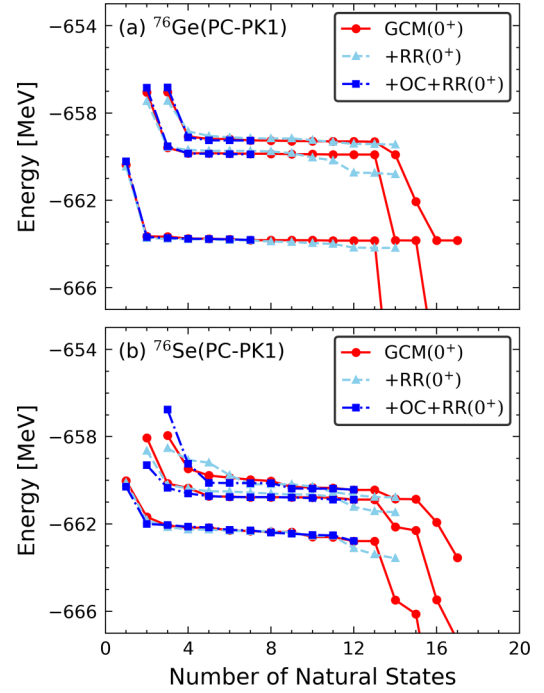

FIG. 13. Same with Fig. 7, but for the relativistic EDF PC-PK1.

OC + RR compared to the full GCM calculation. We note that in the case with sparsely distributed HFB states, the results are somewhat sensitive to the choice of the degree N in the polynomial regression. When a set of denser mesh points with $\Delta\beta = 0.1$ or $\Delta\beta = 0.05$ is employed, the GCM + OC + RR can reproduce the results of the full GCM calculation, and the results are much more robust against variations of the parameter N. In short, the main findings in the results of the relativistic PC-PK1 EDF are generally similar to those of the Gogny D1S force.

B. Hamiltonian-based GCM calculations

In this subsection, we present results for Hamiltonian-based GCM calculations, where the shell-model interaction GCN2850 [85] and the *magic* chiral $NN + 3N$
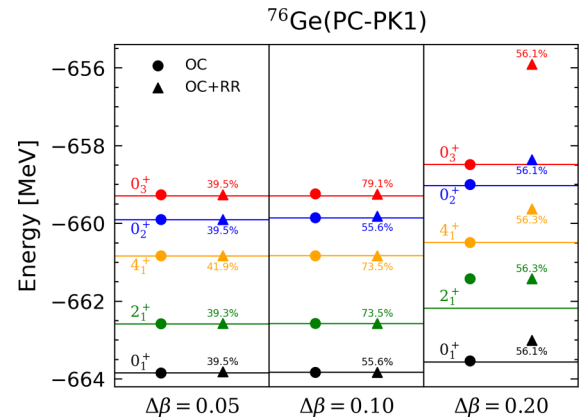

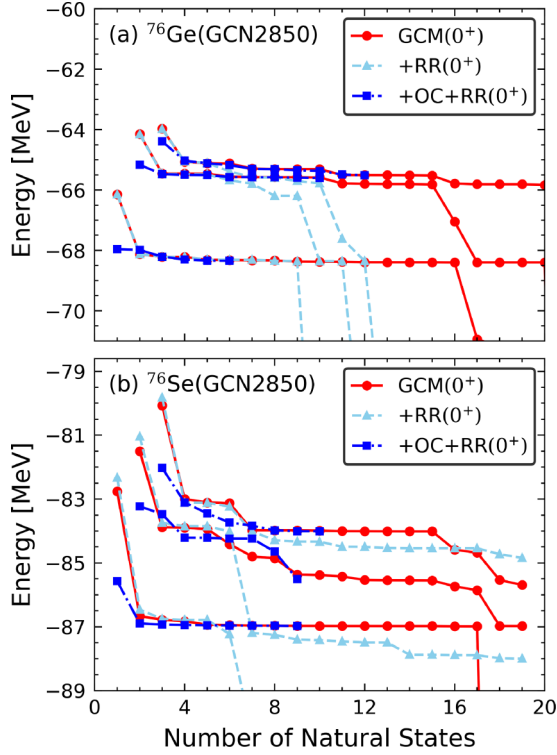FIG. 14. Same as Fig. 11, but for the relativistic EDF PC-PK1.

FIG. 15. Same as Fig. 7, but for the shell-model interaction GCN2850.

FIG. 16. Same as Fig. 7, but for the chiral $2N + 3N$ interaction EM1.8/2.0 with $e_{\mathrm{Max}} = 6$ and $\hbar\omega = 12$ MeV.

interaction EM1.8/2.0 [86,87] are employed. The EM1.8/2.0 interaction provides excellent empirical agreement between ground-state energies and data through at least the $A \sim 60$–70 mass region [88], which is why it has been frequent used in recent *ab initio* studies, including the in-medium GCM studies of light nuclei [89,90] and ^{48}Ca [27] by some of the authors of this work. This interaction, together with the decay operator, is evolved with the multi-reference in-medium similarity renormalization group [91,92] with a reference ensemble comprising prolate, spherical, and oblate HFB minima in both ^{76}Ge and ^{76}Se and with $e_{\mathrm{max}} = 6$. Here, only axially deformed configurations are employed in the GCM calculations. Therefore, the results are somewhat different from those published in Ref. [31], where triaxially deformed configurations and those with neutron-proton pairing correlations are also included.

Figure 15 displays the energies of states from the GCM calculations based on the shell-model interaction GCN2850 for ^{76}Ge and ^{76}Se as a function of the number of natural states. The energy plateaus are much worse than those found in the two EDF-based cases. It is even difficult to determine the energies of the states in the GCM + RR calculations. This is especially true for ^{76}Se. The inclusion of the OC method improves the situation, and the energies of both nuclei by the full GCM calculations are reasonably reproduced by the GCM + OC + RR.

Figure 16 displays the energies of the first three 0^+ states starting from the chiral nuclear force EM1.8/2.0. We find that the energy plateaus in the full GCM calculation are comparable to those found in the EDF-based calculations, and they are
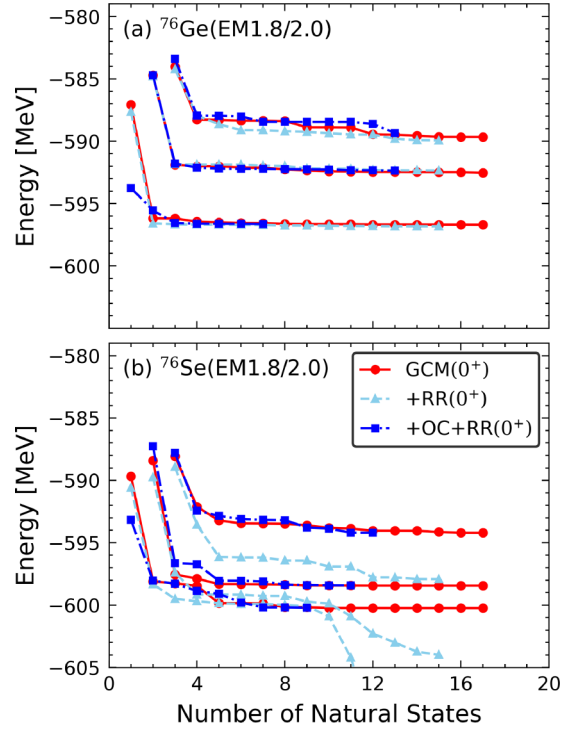
well reproduced in the GCM + OC + RR calculations. The energies of the states in ^{76}Se can hardly be reproduced by the GCM + RR only, demonstrating again the necessity and success of implementing the OC method additionally.

C. Energy spectra and nuclear matrix elements of $0\nu\beta\beta$ decay

Figure 17 shows the data on the low-energy spectra of both ^{76}Ge and ^{76}Se, in comparison with the GCM + OC + RR calculations based on two EDFs and two Hamiltonians. Results of calculations using the configurations tailored for each state and those of calculations using the same configurations as that for the ground state are shown. One can see that with the mixing of axially deformed configurations, the main feature of the sequences $(0_1^+, 2_1^+, 4_1^+)$ in the energy spectra is reasonably reproduced, even though the predicted energy spectra from the calculations with the EDFs (D1S, PC-PK1), and chiral interaction EM1.8/2.0 are more spread out, in contrast to the results of GCN2850. Besides, it is shown that the two different subspace selection strategies yield only slightly different excitation energies for the high-lying states in all the calculations except for GCN2850, where the excited 0^+ states are shifted much higher when restricted to the configurations of the subspace for the ground state. This is probably due to the limited number of valence single-particle states based on which one cannot sample the configurations with large quadrupole deformation.

Figure 18 displays the convergence of the NME for the $0\nu\beta\beta$ decay of ^{76}Ge with respect to the total number of configurations in ^{76}Ge and ^{76}Se from the calculations with different EDFs and interactions for a given cutoff value L_c.
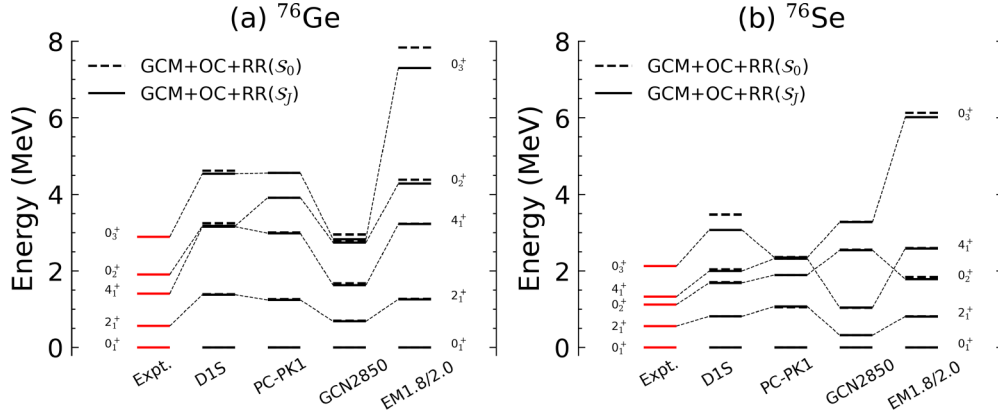
FIG. 17. The low-lying energy spectra of (a) ^{76}Ge and (b) ^{76}Se from GCM + OC + RR calculations based on two different nuclear Hamiltonian and two EDFs. The solid lines are the results of calculations using the configurations within the subspace (S_J) for each state. The dashed lines are the results of calculations for all the states using the configurations in the same subspace (S_0) as that for the ground state. The results are compared to data from Ref. [82].

The L_c is determined in such a way that the convergent value of the NME is rather stable with the increase of the L_c, see Ref. [31]. It is shown in Fig. 18 that the NME in each case converges reasonably well with the increase of the number of states within the subspace, even though the convergence behavior is slightly worse for the PC-PK1 case. Figure 19 summarizes all the NMEs from different calculations. One can see that the NMEs by the GCM calculations in full space can be excellently reproduced in the GCM+OC/ENTROP+RR, with about one order of magnitude reduced computational cost. Again, it should be emphasized that these NMEs cannot be interpreted as the final NME by each EDF or interaction as only axially deformed configurations are considered here.
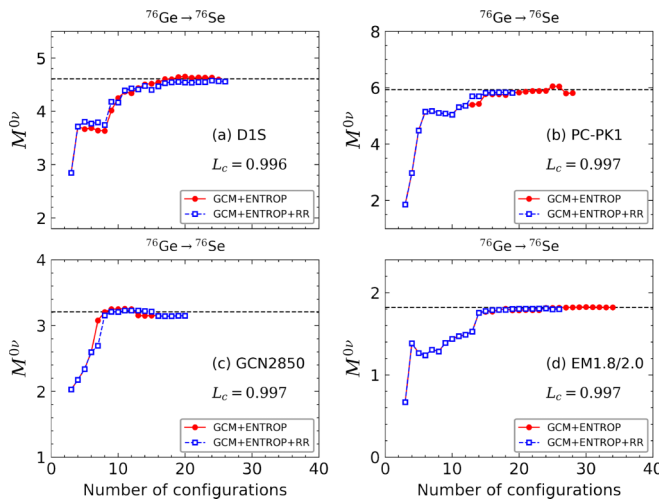
IV. CONCLUDING REMARKS

The representation of wave functions for the nuclear states of interest in terms of a set of nonorthogonal basis functions is the core idea of GCM. In this approach, the nuclear many-body problem is transformed into a generalized eigenvalue problem, where the dimension of the norm and Hamiltonian kernels grows significantly with the number of the collective coordinates in the GCM. Therefore, finding an efficient way to sample the basis functions in the multi-dimensional parameter space is important to accelerate or emulate nuclear model calculations without loss of accuracy.

In this work, we have explored different procedures for implementing statistical ML techniques into GCM calculation to reduce its computational complexity. To mitigate the impact of noise in the predicted kernels by the RR model, we have proposed a subspace-reduction algorithm in which optimal ML models are used as a surrogate method for the exact quantum-number-projection calculation of the norm and Hamiltonian kernels. The efficiency and accuracy of each procedure are demonstrated with both non-relativistic and rel-
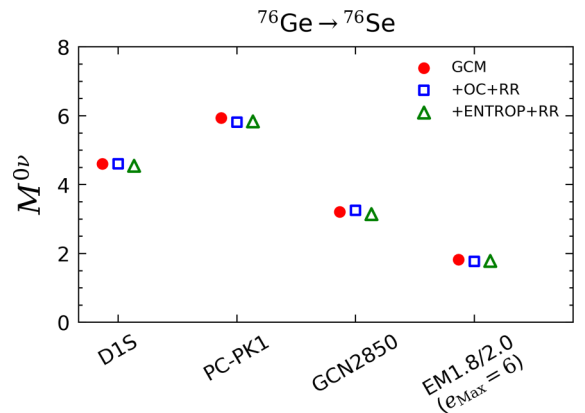

FIG. 18. The convergence of the NME for the $0\nu\beta\beta$ decay of ^{76}Ge against the number of states in ^{76}Ge and ^{76}Se from the GCM+ENTROP (red line) and GCM+ENTROP+RR (blue dotted line) with (a) D1S, (b) PC-PK1, (c) GCN2850, (d) EM1.8/2.0 ($e_{\mathrm{Max}} = 6$), respectively. The results by the original GCM calculation in full space for both nucleus are indicated with dashed lines.

FIG. 19. Comparison of NMEs for the $0\nu\beta\beta$ decay of ^{76}Ge from the full GCM, GCM+OC+RR, and GCM+ENTROP+RR calculations.

ativistic EDFs, a valence-space shell-model Hamiltonian, and a chiral nuclear interaction in calculations for the low-lying energy spectra of ^{76}Ge and ^{76}Se, as well as the $0\nu\beta\beta$-decay NME between their ground states. For the present proof-of-concept study, only axially deformed configurations have been considered. A polynomial RR model was used to learn the norm and Hamiltonian kernels. The results have shown that the noise introduced by the optimal RR model may spoil the description of GCM for nuclear spectra, but this issue can be overcome by applying the subspace-reduction algorithms based on the linear dependence/orthogonality conditions for the basis functions. For the NME of ground-state to ground-state $0\nu\beta\beta$ decay, which requires the wave functions of two different nuclei, we have used both the OC and ENTROP methods to select the subspace, which produce similar results. We have found that in the present study the space-reduction algorithm can speed up the GCM calculation by a factor up to about three to nine for the energy spectra and NME, respectively, with negligible loss in accuracy. One can anticipate that this factor will be significantly larger in GCM calculations with multiple generator coordinates. Extensions of our approach in this direction are in progress.

[1] D. L. Hill and J. A. Wheeler, Phys. Rev. **89**, 1102 (1953).

[2] J. J. Griffin and J. A. Wheeler, Phys. Rev. **108**, 311 (1957).

[3] P. Ring and P. Schuck, *The Nuclear Many-body Problem* (Springer-Verlag, New York, 1980).

[4] P. G. Reinhard and K. Goeke, Rep. Prog. Phys. **50**, 1 (1987).

[5] K. Capelle, J. Chem. Phys. **119**, 1285 (2003).

[6] O. E. Alon, A. I. Streltsov, and L. S. Cederbaum, Phys. Rev. B **71**, 125113 (2005).

[7] E. Orestes, K. Capelle, A. B. F. da Silva, and C. A. Ullrich, J. Chem. Phys. **127**, 124101 (2007).

[8] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Rev. Mod. Phys. **75**, 121 (2003).

[9] T. Nikšić, D. Vretenar, and P. Ring, Prog. Part. Nucl. Phys. **66**, 519 (2011).

[10] J. L. Egido, Phys. Scr. **91**, 073003 (2016).

[11] L. M. Robledo, T. R. Rodríguez, and R. R. Rodríguez-Guzmán, J. Phys. G: Nucl. Part. Phys. **46**, 013001 (2019).

[12] J. A. Sheikh, J. Dobaczewski, P. Ring, L. M. Robledo, and C. Yannouleas, J. Phys. G: Nucl. Part. Phys. **48**, 123001 (2021).

[13] J. M. Yao, K. Hagino, Z. P. Li, J. Meng, and P. Ring, Phys. Rev. C **89**, 054306 (2014).

[14] B. Bally, B. Avez, M. Bender, and P. H. Heenen, Phys. Rev. Lett. **113**, 162501 (2014).

[15] T. R. Rodríguez, Phys. Rev. C **90**, 034306 (2014).

[16] J. L. Egido, M. Borrajo, and T. R. Rodríguez, Phys. Rev. Lett. **116**, 052502 (2016).

[17] J. M. Yao, L. S. Song, K. Hagino, P. Ring, and J. Meng, Phys. Rev. C **91**, 024316 (2015).

[18] E. F. Zhou, J. M. Yao, Z. P. Li, J. Meng, and P. Ring, Phys. Lett. B **753**, 227 (2016).

[19] R. N. Bernard, L. M. Robledo, and T. R. Rodríguez, Phys. Rev. C **93**, 061302(R) (2016).

[20] M. Borrajo and J. L. Egido, Phys. Lett. B **764**, 328 (2017).

[21] T. R. Rodríguez and G. Martínez-Pinedo, Phys. Rev. Lett. **105**, 252503 (2010).

[22] N. L. Vaquero, T. R. Rodríguez, and J. L. Egido, Phys. Rev. Lett. **111**, 142501 (2013).

[23] L. S. Song, J. M. Yao, P. Ring, and J. Meng, Phys. Rev. C **90**, 054309 (2014).

[24] J. M. Yao and J. Engel, Phys. Rev. C **94**, 014306 (2016).

[25] C. F. Jiao, J. Engel, and J. D. Holt, Phys. Rev. C **96**, 054310 (2017).

[26] J. M. Yao, J. Engel, L. J. Wang, C. F. Jiao, and H. Hergert, Phys. Rev. C **98**, 054311 (2018).

[27] J. M. Yao, B. Bally, J. Engel, R. Wirth, T. R. Rodríguez, and H. Hergert, Phys. Rev. Lett. **124**, 232501 (2020).

[28] J. Engel and J. Menéndez, Rep. Prog. Phys. **80**, 046301 (2017).

[29] J. M. Yao, J. Meng, Y. F. Niu, and P. Ring, Prog. Part. Nucl. Phys. **126**, 103965 (2022).

[30] M. Agostini, G. Benato, J. A. Detwiler, J. Menéndez, and F. Vissani, arXiv:2202.01787 (2022).

[31] A. M. Romero, J. M. Yao, B. Bally, T. R. Rodríguez, and J. Engel, Phys. Rev. C **104**, 054317 (2021).

[32] M. Frosini, T. Duguet, J. P. Ebran, B. Bally, T. Mongelli, T. R. Rodríguez, R. Roth, and V. Somà, Eur. Phys. J. A **58**, 63 (2022).

[33] M. Frosini, T. Duguet, J.-P. Ebran, B. Bally, H. Hergert, T. R. Rodríguez, R. Roth, J. M. Yao, and V. Somà, Eur. Phys. J. A **58**, 64 (2022).

[34] T. Duguet, J. P. Ebran, M. Frosini, H. Hergert, and V. Somà, arXiv:2209.03424 (2022).

[35] M. Bender, G. F. Bertsch, and P. H. Heenen, Phys. Rev. Lett. **94**, 102503 (2005).

[36] T. R. Rodríguez, A. Arzhanov, and G. Martínez-Pinedo, Phys. Rev. C **91**, 044315 (2015).

[37] J. Broeckhove and E. Deumens, Z. Phys. A **292**, 243 (1979).

[38] J. Martínez-Larraz and T. R. Rodríguez, Phys. Rev. C **106**, 054301 (2022).

[39] D. Frame, R. He, I. Ipsen, D. Lee, D. Lee, and E. Rrapaj, Phys. Rev. Lett. **121**, 032501 (2018).

[40] A. Sarkar and D. Lee, Phys. Rev. Lett. **126**, 032501 (2021).

[41] A. Ekström and G. Hagen, Phys. Rev. Lett. **123**, 252501 (2019).

[42] S. König, A. Ekström, K. Hebeler, D. Lee, and A. Schwenk, Phys. Lett. B **810**, 135814 (2020).

[43] R. J. Furnstahl, A. J. Garcia, P. J. Millican, and X. Zhang, Phys. Lett. B **809**, 135719 (2020).

[44] C. Drischler, M. Quinonez, P. G. Giuliani, A. E. Lovell, and F. M. Nunes, Phys. Lett. B **823**, 136777 (2021).

[45] D. Bai and Z. Ren, Phys. Rev. C **103**, 014612 (2021).

[46] Y. Kanada-En'yo, Phys. Rev. Lett. **81**, 5291 (1998).

[47] H. Ohta, K. Yabana, and T. Nakatsukasa, Phys. Rev. C **70**, 014301 (2004).

[48] Z.-C. Gao, Phys. Lett. B **824**, 136795 (2022).

[49] T. Otsuka, M. Honma, T. Mizusaki, N. Shimizu, and Y. Utsuno, Prog. Part. Nucl. Phys. **47**, 319 (2001).

[50] S. Shinohara, H. Ohta, T. Nakatsukasa, and K. Yabana, Phys. Rev. C **74**, 054315 (2006).

[51] T. Ichikawa and N. Itagaki, Phys. Rev. C **105**, 024314 (2022).

[52] C. Jiao and C. W. Johnson, Phys. Rev. C **100**, 031303(R) (2019).

[53] D. D. Dao and F. Nowacki, Phys. Rev. C **105**, 054314 (2022).

[54] A. Boehnlein, M. Diefenthaler, N. Sato, M. Schram, V. Ziegler, C. Fanelli, M. Hjorth-Jensen, T. Horn, M. P. Kuchera, D. Lee, W. Nazarewicz, P. Ostroumov, K. Orginos, A. Poon, X.-N. Wang, A. Scheinker, M. S. Smith, and L.-G. Pang, Rev. Mod. Phys. **94**, 031003 (2022).

[55] R. Utama, J. Piekarewicz, and H. B. Prosper, Phys. Rev. C **93**, 014311 (2016).

[56] Z. M. Niu and H. Z. Liang, Phys. Lett. B **778**, 48 (2018).

[57] L. Neufcourt, Y. Cao, W. Nazarewicz, and F. Viens, Phys. Rev. C **98**, 034318 (2018).

[58] Z. M. Niu and H. Z. Liang, Phys. Rev. C **106**, L021303 (2022).

[59] X. H. Wu, Y. Y. Lu, and P. W. Zhao, Phys. Lett. B **834**, 137394 (2022).

[60] R. Utama, W.-C. Chen, and J. Piekarewicz, J. Phys. G: Nucl. Part. Phys. **43**, 114002 (2016).

[61] D. Wu, C. L. Bai, H. Sagawa, and H. Q. Zhang, Phys. Rev. C **102**, 054323 (2020).

[62] X.-X. Dong, R. An, J.-X. Lu, and L.-S. Geng, Phys. Rev. C **105**, 014308 (2022).

[63] N. J. Costiris, E. Mavrommatis, K. A. Gernoth, and J. W. Clark, Phys. Rev. C **80**, 044332 (2009).

[64] Z. M. Niu, H. Z. Liang, B. H. Sun, W. H. Long, and Y. F. Niu, Phys. Rev. C **99**, 064307 (2019).

[65] Z.-A. wang, J. Pei, Y. Liu, and Y. Qiang, Phys. Rev. Lett. **123**, 122501 (2019).

[66] R.-D. Lasseri, D. Regnier, J.-P. Ebran, and A. Penon, Phys. Rev. Lett. **124**, 162502 (2020).

[67] Z.-X. Yang, X.-H. Fan, T. Naito, Z.-M. Niu, Z.-P. Li, and H. Liang, arXiv:2205.15649 (2022).

[68] X. H. Wu, Z. X. Ren, and P. W. Zhao, Phys. Rev. C **105**, L031303 (2022).

[69] S. Islam, H. Mang, and P. Ring, Nucl. Phys. A **326**, 161 (1979).

[70] J. M. Yao, J. Meng, P. Ring, and D. Pena Arteaga, Phys. Rev. C **79**, 044312 (2009).

[71] K. Hagino, G. F. Bertsch, and P.-G. Reinhard, Phys. Rev. C **68**, 024306 (2003).

[72] B. Sabbey, M. Bender, G. F. Bertsch, and P.-H. Heenen, Phys. Rev. C **75**, 044305 (2007).

[73] D. M. Brink and A. Weiguny, Nucl. Phys. A **120**, 59 (1968).

[74] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning - From Theory to Algorithms* (Cambridge University Press, Cambridge, 2014), pp. I–XVI, 1–397.

[75] A. Geron, *Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, Sebastopol, CA, 2017).

[76] R. J. Barlow, *Statistics. A guide to the Use of Statistical Methods in the Physical Sciences (Manchester Physics Series)* (John Wiley & Sons, New York, 1989).

[77] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, Phys. Rep. **810**, 1 (2019).

[78] J. Decharge and D. Gogny, Phys. Rev. C **21**, 1568 (1980).

[79] J. Berger, M. Girod, and D. Gogny, Comput. Phys. Commun. **63**, 365 (1991).

[80] J. M. Yao, arXiv:2204.12126 (2022).

[81] C. Hicks and D. Lee, arXiv:2209.02083 (2022).

[82] National Nuclear Data Center, NuDat2 Database (2020), https://www.nndc.bnl.gov/nudat2.

[83] P. W. Zhao, Z. P. Li, J. M. Yao, and J. Meng, Phys. Rev. C **82**, 054319 (2010).

[84] L. S. Song, J. M. Yao, P. Ring, and J. Meng, Phys. Rev. C **95**, 024305 (2017).

[85] J. Menéndez, A. Poves, E. Caurier, and F. Nowacki, Nucl. Phys. A **818**, 139 (2009).

[86] K. Hebeler, S. K. Bogner, R. J. Furnstahl, A. Nogga, and A. Schwenk, Phys. Rev. C **83**, 031301(R) (2011).

[87] A. Nogga, S. K. Bogner, and A. Schwenk, Phys. Rev. C **70**, 061002(R) (2004).

[88] S. R. Stroberg, J. D. Holt, A. Schwenk, and J. Simonis, Phys. Rev. Lett. **126**, 022501 (2021).

[89] J. M. Yao, A. Belley, R. Wirth, T. Miyagi, C. G. Payne, S. R. Stroberg, H. Hergert, and J. D. Holt, Phys. Rev. C **103**, 014315 (2021).

[90] J. M. Yao, I. Ginnett, A. Belley, T. Miyagi, R. Wirth, S. Bogner, J. Engel, H. Hergert, J. D. Holt, and S. R. Stroberg, Phys. Rev. C **106**, 014315 (2022).

[91] H. Hergert, Phys. Scr. **92**, 023002 (2017).

[92] H. Hergert, S. K. Bogner, T. D. Morris, A. Schwenk, and K. Tsukiyama, Phys. Rep. **621**, 165 (2016).