LUND UNIVERSITY
Faculty of Science

Reaching for the limit of stability

Weiyi Cui

Thesis submitted for the degree of Master of Science
Project duration: 3.5 months

Supervised by Gillis Carlsson and Andrea Idini

Department of Physics
Division of Mathematical Physics
December 2019

Acknowledgements

I want to thank my supervisors Gillis Carlsson and Andrea Idini for their kind helps on my project. They always have wonderful ideas and their patience and kindness ease my stress quite a lot through the whole work. Their passions in work sometimes bring motivations to me as well.

I also want to express my appreciate to my parents. Although they are not around, they always try to encourage me whenever I feel low. I feel lucky to have such a family.

Last but not least, huge thanks to all my friends who shared happiness and sadness with me during this period. I am afraid if I will forget anyone so I decide not to list their names here. However, I am sure you will know when you happen to read this paper. Thank you!

Abstract

In this project, calculations for the total binding energy of all even-even nuclei available from experimental data are performed using the HFBTHO (axially deformed configurational Hartree-Fock-Bogoliubov calculations with Skyrme-forces and zero-range pairing interaction using Harmonic-Oscillator and/or Transformed Harmonic-Oscillator states [1]) program. An artificial neural network is applied to train the information obtained from the HFBTHO calculations and predict the binding energy for the nuclei. The results show impressive improvements to the HFBTHO program. In recent years, the combination of scientific research and machine learning algorithms has become a popular and successful practice. Although it is hard to judge whether an algorithm is good enough, especially with the rapid development of computer science, the application of machine learning in nuclear models can be reliable and promising in predicting the nuclear properties.

List of abbreviations

SCMF	Self-consistent mean-field
EFT	Effective field theory
QCD	Quantum chromodynamics
CI	Configuration interaction
DFT	Density functional theory
HFB	Hartree-Fock-Bogoliubov
BE	Binding energy
ML	Machine learning
NN	Neural network
RMS	Root-mean-square

Contents

Chapter I
Introduction

The topic of discovering new elements remains popular in recent decades. Studies of these new superheavy elements are on the forefront of nuclear, atomic physics and chemistry research. With the capability of the new generation accelerators to detect events with picobarn cross sections, such as the facilities in Superheavy Element Factory (SHEF) in Dubna, new elements are not far from unraveled. The binding energies of the new elements are especially important because one can use the energies to determine the α decay lifetimes and therefore the stability of the new elements. In this project, an energy density functional (EDF) model is combined with a neural network to predict the binding energies for all even-even nuclei, which have even numbers of protons and neutrons, that are experimentally available. The neural network, as a popular approach in modern research, serves as an improvement to the original theoretical model.

In section 1, the theories behind the nuclear program HFBTHO used in this project are illustrated, mainly focusing on the Hartree-Fock-Bogoliubov approach. Section 2 introduces the basic concept of the artificial neural network which is widely used in nuclear investigations. The detailed parameters of the HFBTHO program and the network are explained in chapter II, where for the neural network, the python machine learning library Keras based on Tensorflow [2] is used and the main codes of the network are generated from Idini's work [3]. The investigations based on the HFBTHO program are shown in section 5. Theses results have a root-mean-square deviation of 6.4597MeV over all even-even nuclei and the implementation of the network leads to a lowest deviation of 0.218308MeV in section 6. A final summary and outlook is given in chapter IV.

1 Nuclear physics

1.1 Nuclear models

The nucleus was first discovered by the Geiger-Marsden experiment between 1908 and 1913 [4]. Since then, different nuclear models had been proposed to understand the behavior of nuclei and nucleons inside. One can divide these models into four categories: models based on *ab initio* description; macroscopic models; models based on a self-consistent mean field (SCMF) and shell models [5].

Ab initio methods focus on the given nucleon-nucleon potential while solving the non-relativistic Schrödinger equation. This effective field theory (EFT) favors the application of quantum chromodynamics (QCD) in the low-energy regime of the nuclear interaction [6, 7]. However, a three-body interaction term should be included to give a quantitative description of the nuclear matter rather than only the nucleon-nucleon interaction. The problem of how to implement this term is still under debate.

In macroscopic models, such as the nuclear liquid-drop model [8], global properties of a nucleus are investigated. An energy correction for the quantum shell structure is often applied, leading to the microscopic-macroscopic (mic-mac) method. This method tends to have good results on the calculations of nuclear binding energies, but for unknown exotic nuclei, the mic-mac method can have low reliability.

In the nuclear shell model, one can calculate the properties of a nucleus by constructing the Hamiltonian of the Schrödinger equation with a one-particle operator $\hat{H}_1$ and a two-

particle operator $\hat{H}_2$(1.1) ,

$$\hat{H} = \hat{H}_1 + \hat{H}_2 = \sum_{i=1}^{N}[-\frac{\hbar^2}{2m}\nabla_i^2 + V(r)] + \frac{1}{2}\sum_{j\neq i}^{N} U(\mathbf{r}_i, \mathbf{r}_j), \tag{1.1}$$

where N is the number of nucleons, m is the mass of the nucleon, $\hbar$ is the Planck constant and ∇ is the laplacian operator. $\hat{H}_1$ is constituted of the kinetic energy and mean field potential of the particle contributed by nuclear forces and $\hat{H}_2$ represents the interaction between nucleons. The eigenfunction of this Hamiltonian will be a single Slater determinant, which describes the wavefunctions. To solve a many-body problem, the wavefunction should be expanded into a linear combination of several states which mix with each other to minimize the energy. The calculation then turns into solving a matrix of Slater determinants [9]. This method is called configuration interaction (CI) method and the dimension of the matrix can grow rapidly with more and more nucleons involved.

The SCMF approach in nuclear physics is based on the nuclear density functional theory (DFT) [10]. In the Hohenberg-Kohn theorems, a many-fermion system has a universal energy density functional of the local density distribution [10]:

$$\hat{H}\Psi = [\hat{T} + \hat{V} + \hat{U}]\Psi = [\sum_{i}^{N}(-\frac{\hbar^2}{2m_i}\nabla_i^2) + \sum_{i}^{N} V(\mathbf{r}_i) + \sum_{i<j}^{N} U(\mathbf{r}_i, \mathbf{r}_j)]\Psi = E\Psi, \tag{1.2}$$

where $\hat{T}$, the kinetic energy and $\hat{U}$, the fermion-fermion interaction energy, are the universal operators which keep the same for any N-fermion system. $\hat{V}$ is the potential energy, depending on the certain system. The universal operators $\hat{T}$, $\hat{U}$ and non-universal operator $\hat{V}$ can be written into a functional of the fermion density $n(\mathbf{r})$. Then, the solution of (1.2) becomes the minimization of the functional:

$$E[n] = T[n] + U[n] + \int V(\mathbf{r})n(\mathbf{r})d^3r. \tag{1.3}$$

Therefore, one can determine the nuclear energy density functional $E[n]$ with the properties of nucleons, such as spins, momentum and kinetic energy. Then, the approximate solution to the Schrödinger equation requires calculations of a *self-consistent* mean field. The SCMF approach is successful in nuclear calculations, especially rotational bands in heavy nuclei, and also seems promising in superheavy nuclei with developing corrections to the mean field term. In this project, one focuses specifically on one of the nuclear SCMF models, the Hartree-Fock-Bogoliubov (HFB) method.

1.2 The Hartree-Fock-Bogoliubov method

SCMF models in the particle basis (e.g. Hartree-Fock approach) only take the particle-hole interaction into account, where there is a clear difference between occupied and unoccupied states [11]. This leads to a successful approximation in nuclei with closed shell, but for most nuclei, one also needs to consider the particle-particle or pairing interaction. For these nuclei, the mean field without taking pairing correlations into consideration does not perform well. Therefore, the HFB method is introduced to solve this problem for the mean field.

1.2.1 The Hartree-Fock approach

The Hartree-Fock method was introduced to solve the ground state of many-body problems [12, 13, 14] and was mostly applied in atomic physics at first. In nuclear physics, the problem is formalated in terms of the Schrödinger equation with the Hamiltonian,

$$\hat{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{A}\nabla_i^2 + \sum_{i<j}V(x_i - x_j), \tag{1.4}$$

where $A = N + Z$ is the sum of N, the number of neutrons and Z, the number of protons. V is the given potential and x_i is the generic point in $\mathbb{R}^3$. The wavefunctions are chosen to be

$$\Phi(x_1, \ldots, x_A) = \frac{1}{\sqrt{A!}}\sum_{\sigma}(-1)^{|\sigma|}\prod_{i=1}^{A}\psi_{\sigma(i)}(x_i) = \frac{1}{\sqrt{A!}}det(\psi_i(x_j)), \tag{1.5}$$

where σ is the permutation of the nucleon and $\psi_1, \ldots, \psi_A$ are A functions in $\mathbb{R}^3$, in this case, the wavefunctions of a certain nucleon. This expression is also called a Slater determinant. The final energy obtained from the Hartree-Fock minimization is expressed as [15],

$$E_{HF} = Inf\left\{E(\psi_1, \ldots, \psi_A), \int \psi_i\psi_j^* dx = \delta_{ij} \qquad for \quad 1 \leqslant i, j \leqslant A\right\}, \tag{1.6}$$

where $E(\psi_1, \ldots, \psi_A)$ has the form,

$$E(\psi_1, \ldots, \psi_A) = \frac{\hbar^2}{2m}\sum_{i=1}^{A}\int |\nabla\psi_i|^2 dx + \frac{1}{2}\sum_{i,j}^{A}\int\int |\psi_i(x)|^2 V(x-y)|\psi_j(y)|^2 dxdy$$
$$- \frac{1}{2}\sum_{i,j}^{A}\int\int \psi_i(x)\psi_j^*(x)V(x-y)\psi_i^*(y)\psi_j(y)dxdy. \tag{1.7}$$

In equation (1.7), the second term is called the direct term and the third term is the exchange term. In nuclear physics calculations, the interaction term in $E(\psi_1, \ldots, \psi_A)$ can have slight differences according to what forces are used.

In the expression of the Hartree-Fock minimization, one usually denotes the density of kinetic energy $\sum_{i=1}^{A}|\nabla\psi_i|^2$ by τ, $\sum_{i=1}^{A}|\psi_i(x)|^2$ by ρ and the density matrix $\rho(x, y) = \sum_{i=1}^{A}\psi_i(x)\psi_i^*(y)$. Due to the fact that $\int \psi_i\psi_j^* dx = \delta_{ij}$, one can write the energies as

$$E(\psi_1, \ldots, \psi_A) = \frac{\hbar^2}{2m}\int \tau dx + \frac{1}{2}\int\int \rho(x)V(x-y)\rho(y)dxdy$$
$$- \frac{1}{2}\int\int V(x-y)|\rho(x,y)|^2 dxdy.$$

This shows the energy of a many-body system can be written as a functional of the density, which is the underlay of density-functional theory (DFT).

1.2.2 Pairing correlations

According to the Pauli exclusion principle, each nuclear state can only be occupied by one nucleon or two nucleons with opposite spins. The short-range nuclear force contributes to the mean field potential and the residual interaction between nucleons as well. Therefore, two nucleons can minimize their energy by moving time-reversed orbits due to the residual interaction. For example, in the nuclear shell model, two neutrons or two protons will move into the orbit with the same total angular momentum j but opposite m, where m is the eigenvalue of the z component of j, which means the total angular momentum of the pair is 0.

For nuclei with an even number of neutrons and protons, which is also called even-even nuclei, the ground state is a linear combination of the nucleon pairs coupled in the states near the Fermi energy. The excited states are obtained by breaking the pairs, as figure 1 shows. Including the pairing correlations, quasiparticle states are concerned. The excitation energy is about 2Δ, where Δ is the pairing gap. For nuclei with an odd number of neutrons or protons, the excitation states have an order of Δ with the pairing correlations.

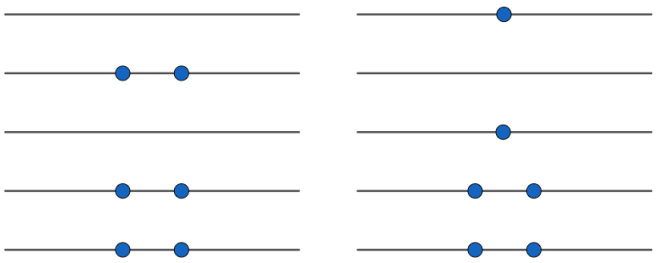


Figure 1: An example of the breaking of a nucleon pair with the extremal single-particle configuration.

The nucleon separation energy is defined as the energy required to separate the last nucleon,

$$S_n(N,Z) = B(N,Z) - B(N-1,Z); \tag{1.8}$$

$$S_p(N,Z) = B(N,Z) - B(N,Z-1). \tag{1.9}$$

$S_n(N,Z)$ is the separation energy for the last neutron and $S_p(N,Z)$ is the separation energy for the last proton. $B(N,Z)$ is the binding energy (BE) of a nucleus with Z protons and N neutrons. Figure 2 shows the experimental neutron separation energies of nuclei with $N-Z = 21, 23$ around the $N = 82$ shell with the data from [16]. The difference in the separation energy of an odd-N nucleus and the even-N nucleus shows one of the evidence of the existence of pairing correlations, which is known as the odd-even staggering. This means it is always more difficult to separate a nucleon from the nucleus when it is in a pair.

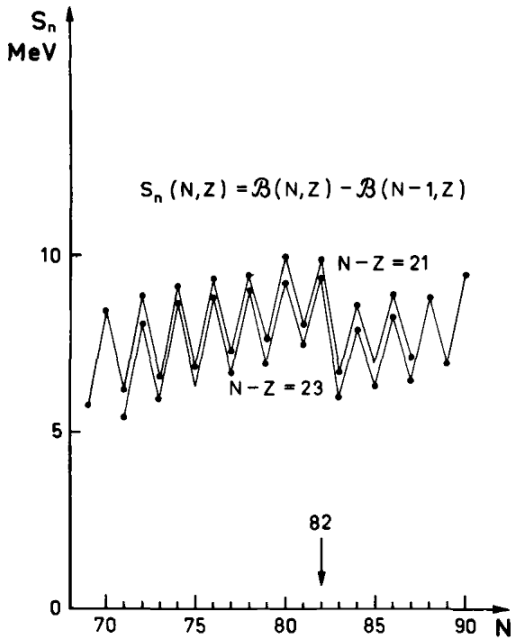


Figure 2: The neutron separation energies [17].

1.2.3 The BCS theory

For the ith nucleon in the single-particle state $\psi_i(\boldsymbol{x})$, one can write the creation operator $\hat{a}_i^+$ as,

$$\hat{a}_i^+ = \int d^3r \sum_{\sigma\tau} \psi_i(\boldsymbol{x}) \hat{a}_x^+, \tag{1.10}$$

where $\hat{a}_x^+$ is the creation operator for the eigenstate at the certain position, and $\boldsymbol{x} = (\boldsymbol{r}, \sigma, \tau)$ with $\sigma = \pm 1$ for spin and $\tau = \pm 1$ for isospin indices [5]. One can then describe the occupied states by $\hat{a}_i^+ |\Phi\rangle$ $(1 \leq i \leq A)$ and unoccupied states $\hat{a}_i |\Phi\rangle = 0$ $(i > A)$, where $|\Phi\rangle = det\{\psi_i(\boldsymbol{x}), i = 1, \ldots, A\}$ is the Slater determinant and A is the number of nucleons.

According to the BCS theory (proposed by Bardeen, Cooper and Schrieffer [18]), nucleons can be treated in pairs moving in time-reversal states, analogous to Cooper pairs in superconductivity which are bound together [19]. The trial wavefunction for a single nucleon in the state μ outside a closed inert core $|0\rangle$ is

$$\Phi_\mu = \prod_{\nu>0, \nu\neq\mu} (U_\nu + V_\nu \hat{a}_{\bar{\nu}}^+ \hat{a}_\nu^+) \hat{a}_\mu^+ |0\rangle \,, \tag{1.11}$$

where $\hat{a}_{\bar{\nu}}^+$ is the creation operator for a nucleon in the time-reversed state $|\bar{\nu}\rangle = (-1)^{j-m} |j - m\rangle$ correlated to the single-particle state $|\nu\rangle = |jm\rangle$. U_ν and V_ν are coefficients that keep the wavefunction normalized. To obtain the best wavefunction, one needs to minimize the expection value $\langle\Phi_\mu| \hat{H} |\Phi_\mu\rangle$, which is equivalent to minimize $\langle\Phi_\mu| \hat{H} - \lambda N |\Phi_\mu\rangle$, where λ is the Fermi energy and N is the number of particles. The Hamiltonian is consisted of a single-particle term and a pairing interaction term,

$$\hat{H} = \sum_{\nu>0} \varepsilon_\nu (\hat{a}_\nu^+ \hat{a}_\nu + \hat{a}_{\bar{\nu}}^+ \hat{a}_{\bar{\nu}}) - \sum_{\nu\nu'>0} G_{\nu\nu'} P_\nu^+ P_{\nu'}, \tag{1.12}$$

where the pair-creation operator $P_\nu^+ = \hat{a}_\nu^+ \hat{a}_{\bar{\nu}}^+$ and the pair-annihilation operator $P_\nu = \hat{a}_\nu \hat{a}_{\bar{\nu}}$. $G_{\nu\nu'}$ is called the pairing strength, which is an amplitude for a nucleon pair to transfer from the state $|\nu'\rangle$ and $|\bar{\nu}'\rangle$ to the other state $|\nu\rangle$ and $|\bar{\nu}\rangle$.

One can simplify the result of the expection value to,

$$\langle\Phi_\mu| H - \lambda N |\Phi_\mu\rangle - \langle\Phi| H - \lambda N |\Phi\rangle = E_\mu = \sqrt{(\varepsilon_\mu - \lambda)^2 + \Delta_\mu^2}, \tag{1.13}$$

where Φ is the wavefunction of the core with even N nucleons, ε_μ is the single-particle energy at the state μ and Δ_μ is the pairing gap. E_μ can also be regarded as the energy needed to add an odd nucleon at the state μ. In equation (1.13), the pairing gap Δ and fermi energy λ are assumed not to change when the extra nucleon is added. However, for a nucleus, the changes in Δ and λ exist indeed. These changes in wavefunctions and excitation energies are called blocking effects.

1.2.4 The Bogoliubov transformation

If one neglects the blocking effects, the Bogoliubov transformation can be introduced for quasiparticles as [11],

$$\begin{aligned} \hat{b}_n^+ &= \sum_i (U_{i,n} \hat{a}_i^+ + V_{i,n} \hat{a}_i), \\ \hat{b}_n &= \sum_i (U_{i,n}^* \hat{a}_i + V_{i,n}^* \hat{a}_i^+), \end{aligned} \tag{1.14}$$

where $\hat{b}_n^+$ represents the creation operator of quasiparticle states and the annihilation operator is $\hat{b}_n$. U and V are the coefficients that transform the single-particle states (i) into quasiparticle states (n). The quasiparticle operators obey the anti-commutation rule,

$$\left\{\hat{b}_\mu^+, \hat{b}_\nu^+\right\} = 0; \left\{\hat{b}_\mu, \hat{b}_\nu\right\} = 0; \left\{\hat{b}_\mu, \hat{b}_\nu^+\right\} = \delta_{\mu\nu}, \tag{1.15}$$

where the subscripts μ and ν refer to the quasiparticle creation or annihilation operator at the state μ and ν. Applying the transformation (1.14) to the rule (1.15), one can get,

$$\begin{aligned} \sum_k (U_{k,\mu} U_{k,\nu}^* + V_{k,\mu} V_{k,\nu}^*) &= \delta_{\mu,\nu} \\ \sum_k (U_{k,\mu} V_{k,\nu} + V_{k,\mu} U_{k,\nu}) &= 0 \end{aligned} . \tag{1.16}$$

The inverse of (1.14) leads to,

$$\begin{aligned} \sum_\mu (U_{j,\mu} U_{k,\mu}^* + V_{j,\mu}^* V_{k,\nu}^*) &= \delta_{k,j} \\ \sum_\mu (U_{j,\mu} V_{k,\mu}^* + V_{j,\mu}^* U_{k,\mu}) &= 0 \end{aligned} . \tag{1.17}$$

Here, j and k are single-particle states. The transformation matrix between single-particle states and quasipartical states can be written as,

$$B_{i,n} \equiv \begin{pmatrix} U_{i,n} & V_{i,n}^* \\ V_{i,n} & U_{i,n}^* \end{pmatrix}, \tag{1.18}$$

and

$$\begin{array}{l} [BB^+]_{k,j} = \delta_{j,k} \\ [B^+B]_{\mu,\nu} = \delta_{\mu,\nu} \end{array} . \tag{1.19}$$

Therefore, in the BCS theory, at the state μ,

$$\begin{array}{l} \hat{b}^+_{\mu} = U_{\mu}\hat{a}^+_{\mu} - V_{\mu}\hat{a}_{\bar{\mu}}, \\ \hat{b}^+_{\bar{\mu}} = U_{\mu}\hat{a}^+_{\bar{\mu}} + V_{\mu}\hat{a}_{\mu}, \\ \hat{b}_{\bar{\mu}} = U_{\mu}\hat{a}_{\bar{\mu}} + V_{\mu}\hat{a}^+_{\mu}, \\ \hat{b}_{\mu} = U_{\mu}\hat{a}_{\mu} - V_{\mu}\hat{a}^+_{\bar{\mu}}. \end{array} \tag{1.20}$$

The ground state $\left|\tilde{O}\right\rangle$ is assumed to be a vacuum for quasiparticle's operators: $\hat{b}_{\mu}\left|\tilde{O}\right\rangle = 0$. A one-quasiparticle state with a quasiparticle in the state μ is $\Phi_{\mu} = \hat{b}^+_{\mu}\Phi$.

1.2.5 The HFB equation

To obtain the ground state of a nucleus, one applies the variational principle which minimizes the energy,

$$E(U,V) = \langle\tilde{O}|\, H\, |\tilde{O}\rangle\,. \tag{1.21}$$

For quasiparticles, one should consider the one-body density matrix ρ and the pairing tensor κ,

$$\rho_{j,k} = \langle\tilde{O}|\, \hat{a}^+_k\hat{a}_j\, |\tilde{O}\rangle = \sum_{\mu} V_{k,\mu}V^*_{j,\mu}, \tag{1.22}$$

$$\kappa_{j,k} = \langle\tilde{O}|\, \hat{a}_j\hat{a}_k\, |\tilde{O}\rangle = \sum_{\mu} U_{j,\mu}V^*_{k,\mu}. \tag{1.23}$$

In coordinate space, the density matrix and pairing tensor are expressed as,

$$\rho(\boldsymbol{x},\boldsymbol{x}') = \sum_n \psi^{(V)}_n(\boldsymbol{x})\psi^{(V)*}_n(\boldsymbol{x}'), \tag{1.24}$$

$$\kappa(\boldsymbol{x},\boldsymbol{x}') = \sum_n \psi^{(U)}_n(\boldsymbol{x})\psi^{(V)*}_n(\boldsymbol{x}'), \tag{1.25}$$

where

$$\begin{pmatrix} \psi^{(V)}_n(\boldsymbol{x}) \\ \psi^{(U)}_n(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} \sum\limits_i V_{i,n}\psi_i(\boldsymbol{x}) \\ \sum\limits_i U_{i,n}\psi_i(\boldsymbol{x}) \end{pmatrix} \tag{1.26}$$

are quasiparticle wavefunctions in coordinate space.

To satisfy (1.16) and (1.17), the generalized density matrix is defined as

$$R = \begin{pmatrix} \rho & -\kappa \\ \kappa^* & I-\rho^* \end{pmatrix}, \tag{1.27}$$

where I is the identity matrix. By introducing the Lagrange parameters for the constraint on the neutron and proton number and considering the unitary condition $R^2 = R$ [20], the Bogoliubov Hamiltonian is denoted as

$$H = \begin{pmatrix} e & \Delta \\ -\Delta^* & -e^* \end{pmatrix}, \tag{1.28}$$

where

$$\begin{matrix} e_{j,k} = h_{j,k} - \lambda_{q_j}\delta_{j,k}, \quad h_{j,k} = \frac{\delta E}{\delta \rho_{k,j}} = h^*_{k,j} \\ \Delta_{j,k} = \frac{\delta E}{\delta \kappa^*_{j,k}} = -\Delta_{k,j} \end{matrix} . \tag{1.29}$$

Here, λ_{q_j} is the Lagrange parameter for the constraint on the neutron or proton number.

The density matrix can be decomposed into spin-isospin terms:

$$\rho(\boldsymbol{x}, \boldsymbol{x}') = \rho(\boldsymbol{r}\sigma\tau, \boldsymbol{r}'\sigma'\tau'). \tag{1.30}$$

σ and τ are spin and isospin states as mentioned in section 1.2.3.

If one includes Skyrme's effective interaction [21] in the HFB method, the local energy density functional will be constituted by two components: the mean-field and pairing energy densities,

$$\mathcal{H}(\mathbf{r}) = H(\mathbf{r}) + \tilde{H}(\mathbf{r}), \tag{1.31}$$

where $H(\mathbf{r})$ and $\tilde{H}(\mathbf{r})$ depend on the following nuclear properties: local particle density $\rho(\mathbf{r})$, local pairing density $\tilde{\rho}(\mathbf{r})$, kinetic energy density $\tau(\mathbf{r})$ and spin-current density $\mathbf{J}_{ij}(\mathbf{r})$. In the HFBTHO program used in this project, one only considers the spin-dependent one-body density matrices, which leads to the time-even pairing density matrix $\tilde{\rho}(\boldsymbol{r}\sigma, \boldsymbol{r}'\sigma') = -2\sigma'\kappa(\boldsymbol{r}, \sigma, \boldsymbol{r}', -\sigma')$ to replace the time-odd pairing tensor κ. Finally, the Skyrme HFB equations have the form of

$$\sum_{\sigma'} \begin{pmatrix} h(\mathbf{r}, \sigma, \sigma') & \tilde{h}(\mathbf{r}, \sigma, \sigma') \\ \tilde{h}(\mathbf{r}, \sigma, \sigma') & -h(\mathbf{r}, \sigma, \sigma') \end{pmatrix} \begin{pmatrix} U(E, \boldsymbol{r}\sigma') \\ V(E, \boldsymbol{r}\sigma') \end{pmatrix} = \begin{pmatrix} E + \lambda & 0 \\ 0 & E - \lambda \end{pmatrix} \begin{pmatrix} U(E, \boldsymbol{r}\sigma) \\ V(E, \boldsymbol{r}\sigma) \end{pmatrix}. \tag{1.32}$$

In simple words, equation (1.32) presents a clear form on how the HFB energy depends on the original mean-field and further nucleon interactions. For more details in the components and properties of the Skyrme force, one can refer to the HFBTHO program description [1] and a review on SCMF models [5].

1.2.6 Solution to the deformed nuclei

For spherical nuclei, a mean field is already enough to describe the nuclear properties with spherical symmetry. The Skyrme HFB equations also give the best performance in the calculations for these spherical nuclei in coordinate space [22]. However, an incomplete filling of a shell can lead to nuclear deformation. The deformation parameter is phenomenologically defined as,

$$\beta_l = \frac{4\pi}{3AR_0^l} \left\langle r^l Y_{l0} \right\rangle, \tag{1.33}$$

where R_0 is the nucleus radius: $R_0 = 1.2A^{1/3}$ and Y_{l0} is a spherical harmonic. In this case, a deformation should also be applied to the mean-field potential to describe the nuclei better. In 1973, Vautherin [23] proposed a method to calculate axially deformed nuclei, which solves the deformed HFB equation by diagonalizing the HFB Hamiltonian

in the configurational space of wavefunctions with the application of certain symmetry and determines the potentials and densities in coordinate space. To shorten the computation time of the nuclear calculations, a restriction on axially-symmetric deformation is imposed in the HFBTHO program. Then, one can use the standard cylindrical coordinates $\mathbf{r} = (r\cos\varphi, r\sin\varphi, z)$ with z as the symmetry axis and the quasiparticle states can be expressed by,

$$\begin{pmatrix} U_k(r,\sigma,\tau) \\ V_k(r,\sigma,\tau) \end{pmatrix} = \chi_{q_k}(\tau)\left[\begin{pmatrix} U_k^+(r,z) \\ V_k^+(r,z) \end{pmatrix} e^{i\Lambda^-\varphi}\chi_{+1/2}(\sigma) + \begin{pmatrix} U_k^-(r,z) \\ V_k^-(r,z) \end{pmatrix} e^{i\Lambda^+\varphi}\chi_{-1/2}(\sigma)\right], \tag{1.34}$$

where the third component J_z of the total angular momentum is a good quantum number and has the eigenvalue Ω_k for the kth state and $\Lambda^\pm = \Omega_k \pm 1/2$. The quasiparticle states are also assumed to be the eigenstates of the third component τ_z of the isospin operator and the eigenvalues are $q_k = -\frac{1}{2}$ for neutrons and $q_k = +\frac{1}{2}$ for protons. If one substitutes equation (1.34) into equation (1.32), only r and z will be involved in the final equations and local densities [23]. Furthermore, for kth state and its time-reversed state $\bar{k}$, their contributions to the densities are identical. If one restricts the summations to positive Ω_k, total results will be multiplied by a factor of 2.

2 Machine learning

2.1 Artificial neural network

The HFB approach considers pairing correlations on the basis of the SCMF method, which already complicates the Hamiltonian to a large content. However, there still exists some deviations between the theoretical results and experimental data, which is inadequate for the study of exotic nuclei, especially when one wants to extrapolate the model to the neutron drip line. The drip line is a boundary with extremal ratio of the proton number to the neutron number, where no more protons or neutrons can be added into the nucleus. For example, in [24, 25] the deviations of different nuclear approaches are discussed. One way to improve the existing models even further is to introduce machine learning (ML) algorithms. A typical ML algorithm contains a set of "samples" as the input parameters with a certain number of "features" or "attributes". For example, if one wants to train the ML codes to predict the total BE of all even-even Pb isotopes with only the neutron and proton numbers as the input, the 22 even-even isotopes form 22 samples and the features include N and Z.

The input set, output set and training process are assigned in different layers. These layered presentations are learned via *neural networks* (NN). One can define a number of neurons in each layer. The neurons are interconnected, depending on the input values, weights and biases, and they generate new values for the next layer together. A typical NN is shown in figure 3, where d_{ji} are weights for the input neutrons to be transformed into hidden neurons, b_j are weights for the hidden neurons to be transformed into the output neuron and c_j and a are biases for the input and output.

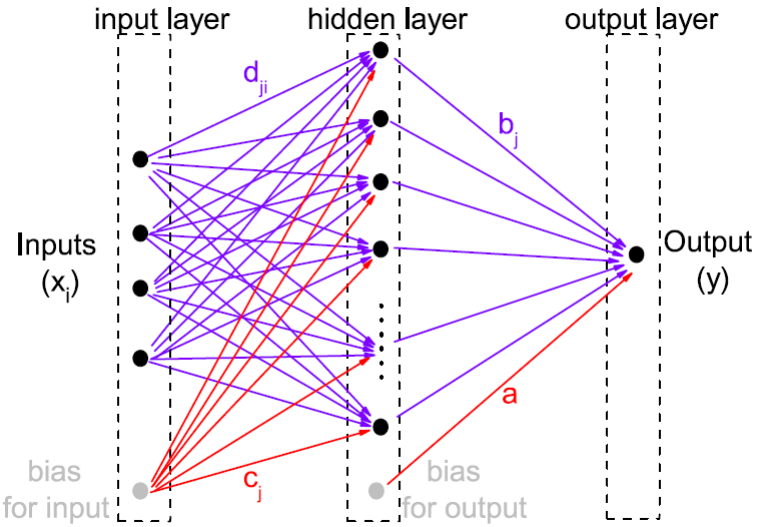


Figure 3: A typical NN with 4 input neurons, 1 bias input neuron and 1 output neuron [26].

For one layer in an artificial NN, the output y can be achieved by the following expression with the input vector $\mathbf{x}(x_1, x_2, \ldots, x_n)$,

$$y = a + \sum_{j=1}^{M} b_j \phi(c_j + \sum_i d_{ji} x_i), \tag{2.1}$$

where M is the number of hidden neurons in one hidden layer and ϕ is a nonlinear activation function to allow a neuron to transfer from an inactivated low value to an activated high value. a and c_j are biases and b_j, d_{ji} are the weights for the corresponding layer. If more hidden layers are involved, the outputs from the former hidden layer will serve as inputs to the next hidden layer, as shown in figure 4.

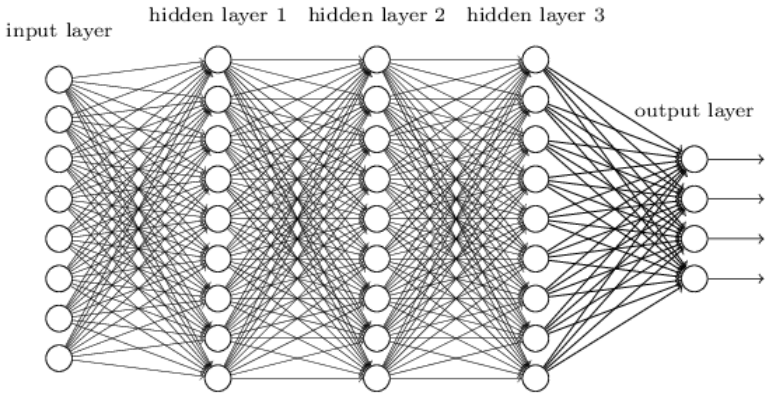


Figure 4: A typical deep NN with 8 input neurons, 4 output neurons and 3 hidden layers [27].

For a layer, a dropout rate can be added, which is an important regularization technique. Dropout represents randomly setting a number of output features of the layer to zero during the training process, while a dropout rate is the fraction of the zeroed features.

The objective of an training process is to minimize a loss function, which computes a distance score between the predictions given by the NN and true targets,

$$Loss = \sum_{i=1}^{n} (y_i - t_i)^2, \tag{2.2}$$

where the output $\mathbf{y}$ is now a vector with n points and t refers to the target value. In some networks, a process called backpropagation also participates in the determination of weights. After one iteration, the NN will propagate the loss back to adjust the weights to minimize the loss function again. Figure 5 shows the summary of how a complete NN works in a flow chart.

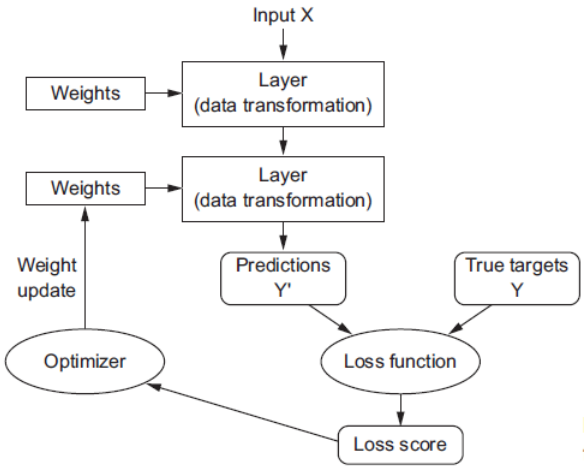


Figure 5: A flow chart of NN mechanism [28].

In recent years, various ML methods are applied not only in software engineering, but also in scientific research projects. One of algorithms, so-called deep learning, is practical and still promising. The "deep" term in the name of deep learning refers to the model with successive layers constructed, which means more types of NN architecture and more hidden layers are included. A more systematical introduction to deep learning can be found in [28].

2.2 Cross validation

Usually when one runs a specific NN, besides the training set, a validation set should also be included. This validation process is important for a model to show whether it still performs well with unknown inputs. However, for the NN used in this project, nuclei in the training are limited to the even-even ones and the number of these nuclei is fixed. Therefore, it is not realistic to find enough "unknown" nuclei for the validation. One has to use another technique in ML called cross validation to solve this problem.

The mechanism of cross validation is quite simple. The main idea is to make full use of the training set itself. One can divide the set into e.g. 5 subsets and in each model-fitting process, one of the subsets is used as the validation set and the rest 4 subsets are used as the training set together. As a result, 5 NN models can be generated with 5 totally different validation sets. The mission is to find the best validation deviation among the 5 runs and save the corresponding model as the best model to fit the data set. Figure 6 visualizes the cross validation method with 5 splits.

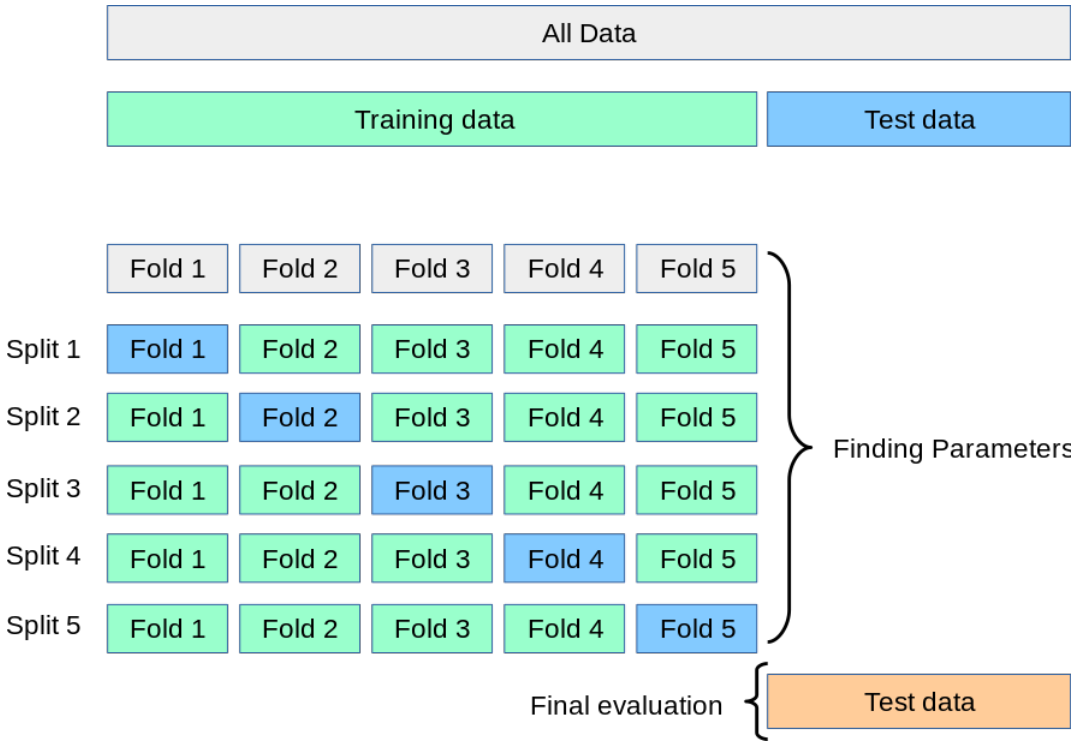


Figure 6: An example of how cross validation works [29].

Chapter II
Method

3 HFBTHO program

The HFBTHO program is named after axially deformed configurational Hartree-Fock-Bogoliubov calculations with Skyrme-forces and zero-range pairing interaction using Harmonic-Oscillator and/or Transformed Harmonic-Oscillator states [1]. To apply the HFB theory into solving the nuclear many-body problem numerically, there are usually three ways. The first method is focused on the diagonalization of the particle-particle part in the Hamiltonian, which is called two-basis method [30, 31, 32]. The second one is called canonical-basis HFB method, which uses the spatially localized eigenstates of the one-body density matrix and the last one is an approach to axial coordinate-space HFB using a basis-spline method [1]. However, the two-basis method can result in a large number of positive-energy free-particle states and the other two consume too much time and computational capability. In the HFBTHO program, Stoitsov *et al.* proposed a configuration-space approach to the deformed HFB equations instead of the coordinate-space approach. In other words, the HFB solution is expanded in the basis of a harmonic oscillator (HO) or a transformed harmonic oscillator (THO). An HO basis is constituted by eigenfunctions of a single-particle Hamiltonian for an axially deformed harmonic oscillator potential [1]. The number of states in the HO basis equals to $(N_{sh}+1)(N_{sh}+2)(N_{sh}+3)/6$, and for a spherical basis, all the shells with N=$0\ldots N_{sh}$ are included, where N_{sh} is the number of shells. Codes based on HFB+HO are successful in most nuclear calculations, but near the nuclear drip lines, they tend to converge slowly, needing a large basis to calculate loosely bound states. One can achieve the THO set where the quasiparticle HFB wavefunctions are expanded by certain means of coordinate transformation. HFB+THO can serve as an alternative to the HFB+HO approach and there are various choices of the THO basis [33, 34]. However, the projection into the THO basis is not used for this project.

In this project, the 2.00d version of the code HFBTHO [35] is used to calculate the binding energies and single-particle states under different quadrupole constraints of even-even nuclei. The energies of the single-particle states are calculated by diagonalizing the mean-field Hamiltonian, and the overlap between the single-particle wave function and the quasiparticle wave function indicates the direction of the single-particle state. Figure 7 shows a typical input file for the HFBTHO program calculating the nucleus $^{138}_{82}$Pb. The number of shells can be varied by the user; the heavier the target nucleus is, the more shells are needed to make the calculation converge. A spherical HO basis is used in this project, and the oscillator length b_0 is determined by the HO frequency $b_0 = \sqrt{\hbar/m\omega_0}$, where $\hbar\omega_0 = 1.2 \times 41/A^{1/3}$ is already defined in the code by default. The accuracy also has a huge impact on the speed of HFBTHO calculations. It is defined as the difference between the result from the current iteration and that from the last iteration. A lower accuracy means a faster convergence, which needs the user to consider the equilibrium in between. In this project, an accuracy of 10^{-6} is used. One can also choose the Skyrme functional to be used in the code, which in this case, is SLy4. Details of the parameters used in SLy4 can be found in [1]. The pairing correlations are also included which add a small number to the pairing matrix elements initially. The Coulomb type chosen to be 2 means both the direct and exchange Coulomb potentials are included. Quardrupole moment constraints are imposed in this project, with the the expectation value of the constraint

calculated by equation (1.33). The finite-temperature HFB equations are also included in the code, where the mean field is described as temperature-dependent [36]. The selection of the input temperature can also affect the converging speed. The investigation into the influence of the number of shells and temperature on the time of calculations will be shown in section 5.1.

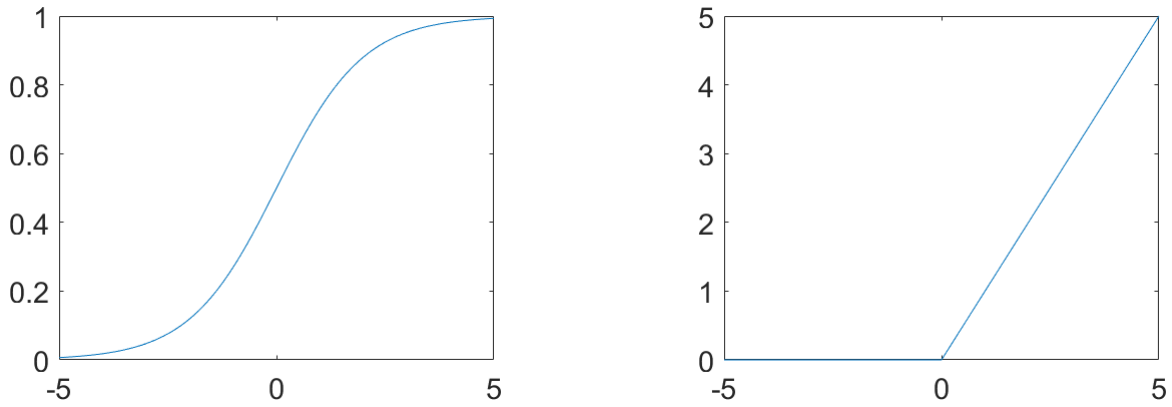
```
&HFBTHO_FILE
 OUTFILE = "out9.out" /
&HFBTHO_GENERAL
 number_of_shells = 14,
 oscillator_length = -1.0,
 basis_deformation = 0.00,
 proton_number = 82, neutron_number = 138, type_of_calculation = 1 /
&HFBTHO_ITERATIONS
 number_iterations = 400, accuracy = 1.E-6, restart_file = -1 /
&HFBTHO_FUNCTIONAL
 functional = 'SLY4', add_initial_pairing = T,
 type_of_coulomb = 2/
&HFBTHO_PAIRING
 user_pairing = F, vpair_n = 0.1, vpair_p = 0.1,
 pairing_cutoff = 60, pairing_feature = 1.0 /
&HFBTHO_CONSTRAINTS
 lambda_values = 1, 2, 3, 4, 5, 6, 7, 8,
 lambda_active = 0, 1, 0, 0, 0, 0, 0, 0,
expectation_values = 0.0, -4.8285803918, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0 /
&HFBTHO_BLOCKING
 proton_blocking = 0, 0, 0, 0, 0, neutron_blocking = 0, 0, 0, 0, 0 /
&HFBTHO_PROJECTION
 switch_to_THO = 0, projection_is_on = 0,
 gauge_points = 1, delta_Z = 0, delta_N = 0 /
&HFBTHO_TEMPERATURE
 set_temperature = T, temperature = 0.04 /
&HFBTHO_DEBUG
 number_Gauss = 40, number_Laguerre = 40, number_Legendre = 80,
 compatibility_HFODD = F, number_states = 1500, force_parity = T,
 print_time = 0 /
```

Figure 7: Typical input parameters for the HFBTHO program.

4 NN configuration

The library Keras based on Tensorflow [2]is used to build the NN for nuclear BE prediction. One can usually customize the configuration of a certain NN to achieve the best performance. The configuration parameters of the NN in this project include:

- number of input and output layers, which is taken as 50 and 1 respectively ;
- number of hidden layers and neurons in each hidden layer, which is chosen to be 2 and 650 for the training on all 859 even-even nuclei experimentally measured (The range of the selection is based on the investigation by Idini [3]);
- epochs: number of iterations performed to minimize the loss function;
- activation functions for input and hidden layers, which are chosen to be “sigmoid” and “ReLU” respectively, shown in figure 8;



(a) Sigmoid (b)Rectified Linear Unit (ReLU)

Figure 8: Different shapes of the activation functions used.

- dropout rate, 0 chosen for this project since there are only 2 hidden layers;
- initialization of weights and biases, which is chosen as a random number from the normal distribution with a mean value of 0 and standard deviation of 0.05;
- optimizer: rmsprop, to speed up the training and reduce the loss by making use of the gradient of the batch;
- size of the cross validation set and number of validation times: 1/10 of the input samples, performed 10 times. A Python library called scikit-learn [29] is used for the data processing in cross validation.

Root-mean-square (RMS) deviation is calculated to evaluate the theoretical or computational total BE of the nuclei (y_i) with the experimental data (t_i):

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(y_i - t_i)^2}{n}}, \tag{4.1}$$

where n is the number of samples.

Chapter III
Results

5 Investigations on the HFBTHO program

5.1 Selection of the input parameters

To perform the calculations on a large range of even-even nuclei, one needs to investigate under which circumstances the codes perform both good and fast calculations. A lighter nucleus ^{48}Ca and a heavier one ^{200}Hg are chosen. Among the input parameters, the number of shells (N_{sh}) and the temperature (T) imposed in the calculations play a crucial role on the performance of the program.

Figure 9 shows how the total binding energies of the two nuclei vary with the number of shells and the temperature. N_{sh} is selected in the range from 8 to 20 for ^{48}Ca and from 10 to 20 for ^{200}Hg. T is changed from 0 to 0.1MeV with an interval of 0.01MeV at $N_{sh} = 10$ for ^{48}Ca and $N_{sh} = 16$ for ^{200}Hg. The total BE continues to decrease rapidly until around 16 shells for both nuclei. For ^{48}Ca, the total BE does not change with the temperature in the range of [0, 0.1MeV] and for ^{200}Hg, the total BE begins to rise when the temperature reaches about 0.04MeV.

Figure 10 shows how the average time for the HFBTHO calculations varies with the number of shells and the temperature. When the number of shells increases, the time increases almost exponentially. However, when a certain temperature is added, the calculation will converge much faster for both nuclei.

The difference between the absolute value of the lowest total BE from the HFBTHO program and from atomic mass evaluation 2016 (AME16) [37, 38] ($|\Delta E|$) and average calculation time (t) with different configurations of N_{sh} and T are listed in Table 1.

For light nuclei such as ^{48}Ca, lower N_{sh} brings both a better theoretical result on the total BE and less time consumed. Larger temperatures lead to smaller t and don't change the value of BE. However, for heavier nuclei like ^{200}Hg, the deviation between theoretical and experimental total BE is too high for a low N_{sh}. Thus, one has to increase N_{sh} to achieve lower $|\Delta E|$, which slows down the convergence in the calculation to a large degree. An increment on T can accelerate the convergence, but will also make $|\Delta E|$ larger for ^{200}Hg. Since the energy difference is much larger in ^{200}Hg compared to that in ^{48}Ca, one should pay more attention to improve the results from heavy nuclei. This means a relatively larger N_{sh} and a moderate T should be chosen in the range of the investigations on ^{48}Ca and ^{200}Hg.

Therefore, the conclusion is that for a balance in the quality and the speed of the calculations based on the HFBTHO codes, one should choose the number of shells around 14 and a temperature less than about 0.04MeV as the input parameters.

	N_{sh}	T/MeV	$\|\Delta E\|$/MeV	t/min
^{48}Ca	8	0.05	0.027847	0.666333
	9	0.05	0.192279	0.71781
	10	0	0.873151	2.76348
	10	0.01	0.873151	1.01567
	10	0.02	0.873151	0.905238
	10	0.03	0.873151	0.86619
	10	0.04	0.873151	0.864524
	10	0.05	0.873151	0.845571
	10	0.06	0.873151	0.839714
	10	0.07	0.873151	0.843381
	10	0.08	0.873151	0.836095
	10	0.09	0.873151	0.833524
	10	0.1	0.873151	0.839857
	11	0.05	1.471553	1.02719
	12	0.05	1.584425	1.19033
	13	0.05	1.647911	1.4721
	14	0.05	1.734281	1.75957
	15	0.05	1.756517	2.20152
	16	0.05	1.855457	2.55305
	17	0.05	1.887971	3.72738
	18	0.05	1.894493	4.73486
	19	0.05	1.897041	6.30729
	20	0.05	1.897293	9.87405
^{200}Hg	10	0.05	14.95328	0.982238
	11	0.05	13.756422	1.20771
	12	0.05	12.403921	1.4581
	13	0.05	10.174596	1.7801
	14	0.05	9.084211	2.23167
	15	0.05	8.836724	3.10167
	16	0	8.558827	7.24352
	16	0.01	8.558827	4.29024
	16	0.02	8.558828	4.07714
	16	0.03	8.558923	4.14314
	16	0.04	8.559893	4.21233
	16	0.05	8.563437	4.07281
	16	0.06	8.571241	3.81148
	16	0.07	8.584285	3.96457
	16	0.08	8.602787	3.93971
	16	0.09	8.626407	3.82143
	16	0.1	8.654492	3.69833
	17	0.05	8.270328	4.82052
	18	0.05	8.112228	5.95905
	19	0.05	7.965679	8.25238
	20	0.05	7.754098	10.6089

Table 1: $|\Delta E|$ and t with different configurations of N_{sh} and T for ^{48}Ca and ^{200}Hg. The green cells show the best result and related parameter in each investigation and the limitation with this configuration is shown in the red cells.

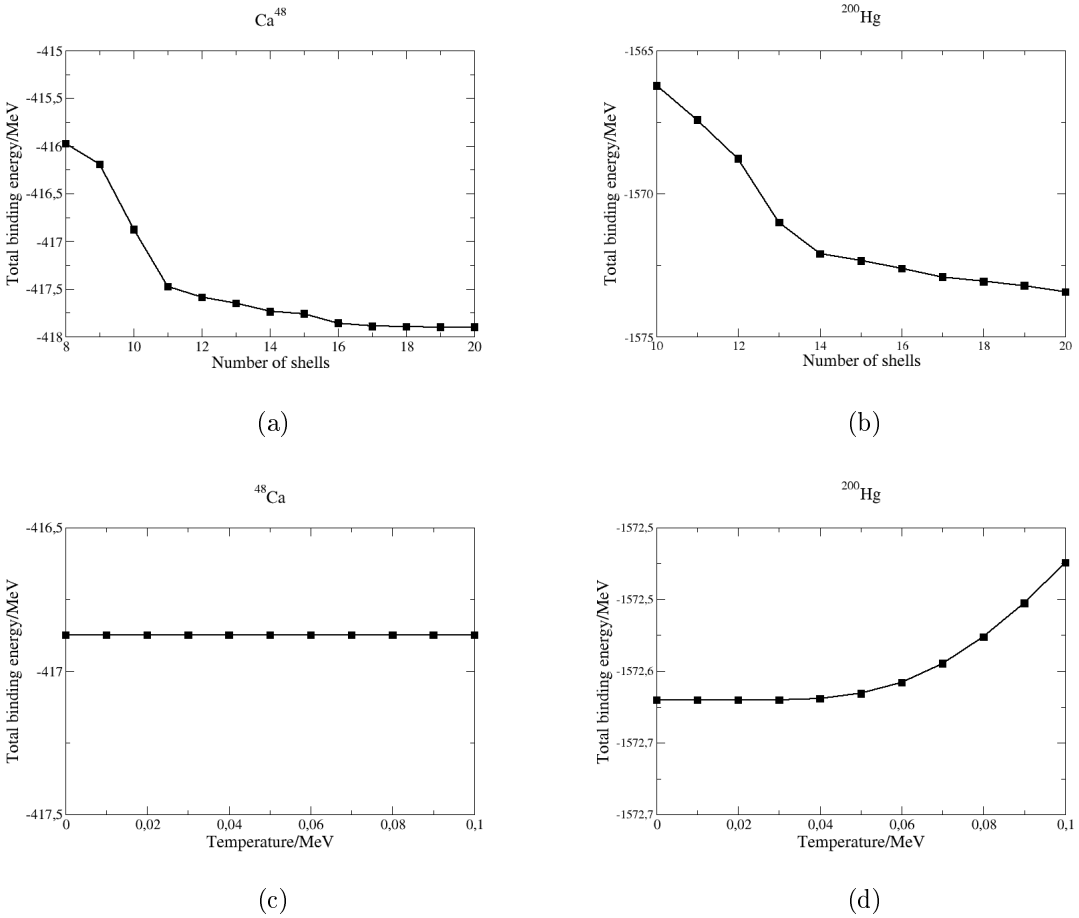


Figure 9: The total binding energies from HFBTHO calculations with a temperature of 0.05MeV and different numbers of shells of: (a) ^{48}Ca; (b) ^{200}Hg, and the total binding energies from HFBTHO calculations with different temperatures: (c) ^{48}Ca with 10 shells; (d) ^{200}Hg with 16 shells.

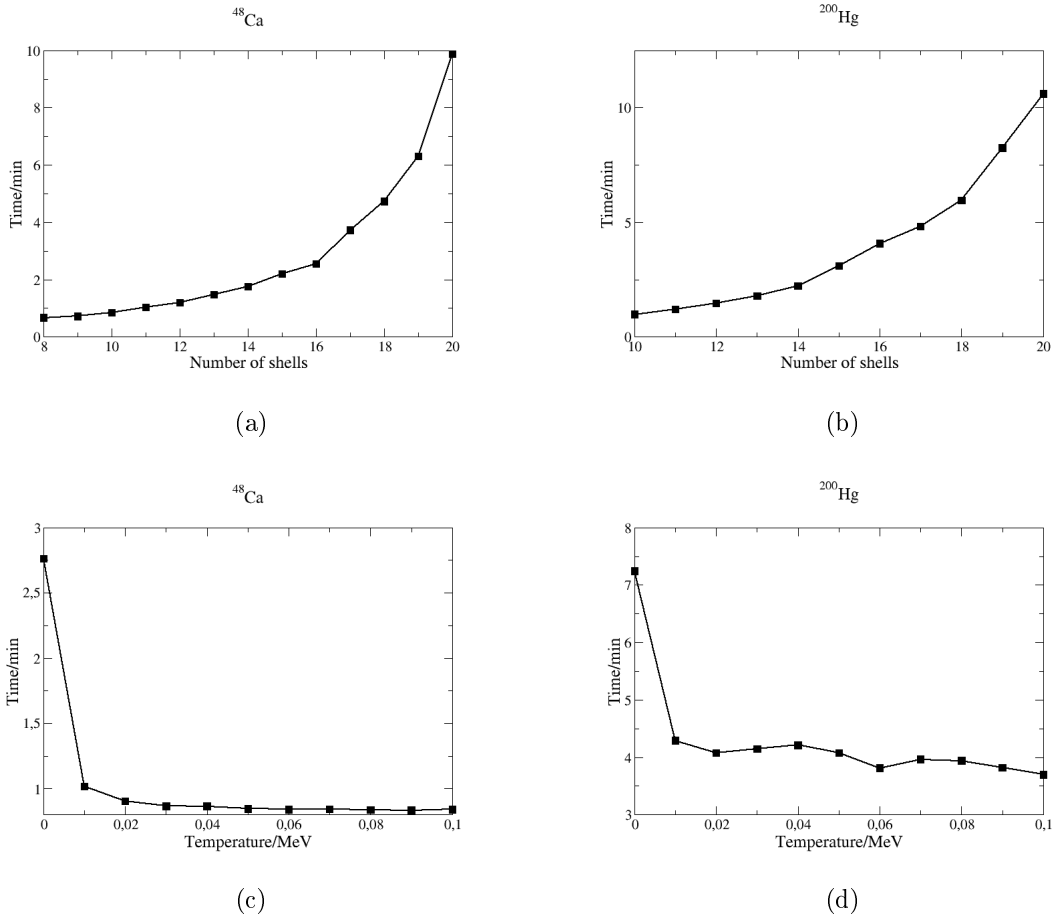


Figure 10: The total CPU times of the HFBTHO calculations with a temperature of 0.05MeV and different numbers of shells of: (a) ^{48}Ca; (b) ^{200}Hg, and the total CPU times of the HFBTHO calculations with different temperatures of: (c) ^{48}Ca with 10 shells; (d) ^{200}Hg with 16 shells.

5.2 Calculations on $_{82}$Pb isotopes

With the parameters concluded from section 5.1, all even-even Pb isotopes are calculated by the HFBTHO program. The number of the even-even isotopes is 22 and 21 different deformations for each nucleus are considered. Figure 11 shows the total BE vs β spectra of these nulcei. The red line in figure 11 represents the total BE spectrum of the doubly magic nucleus $^{126}_{82}$Pb. When the nucleon number gets closer to the magic number, the total BE spectrum becomes more curved. This means the β and BE with different deformations from HFBTHO codes can be good parameters in the NN training, which indicates the information of relatively big energy gaps in the region around $^{126}_{82}$Pb.

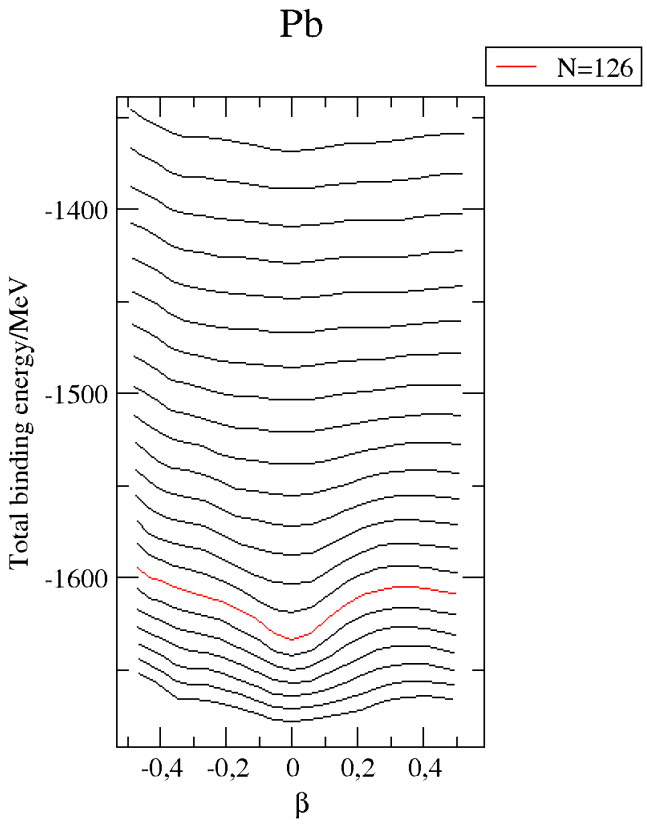


Figure 11: Total binding energies with different deformations of all even-even Pb nuclei.

Figure 12 shows the difference between the lowest total BE calculated by HFBTHO and the BE from AME16 for all Pb even-even nuclei. It also presents a local minima at $^{126}_{82}$Pb. This quantity can also serve as the target value in a NN.

Similar phenomenon happen to the spectra of the neutron pairing energy and the lowest neutron quasiparticle state when BE is the lowest, shown in figure 13 and 14. However, for Pb isotopes, the proton number stays the same so the proton pairing energy in figure 13 does not change either.

All the outputs from HFBTHO discussed in this section can be used as the features in the final NN part because they all show the properties of the nuclei and the trend on how the BE varies to some extents. Then, the investigation can be made on how many features should be fed to the network to get the best result.

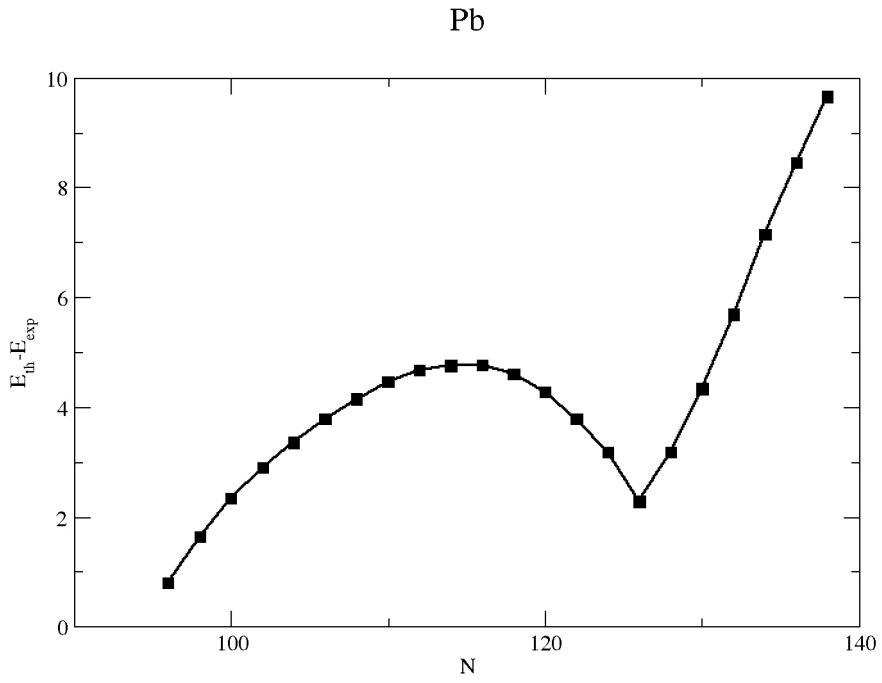


Figure 12: The difference between HFBTHO calculations and AME16 data for even-even Pb nulcei. A local minimum can be observed at $N = 126$.

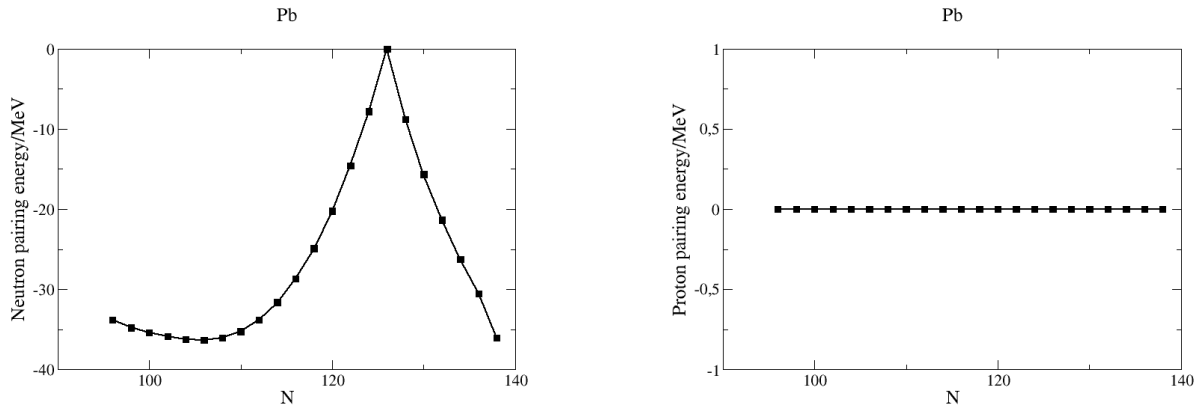


Figure 13: Pairing energy calculated by HFBTHO for neutrons (left) and protons (right).

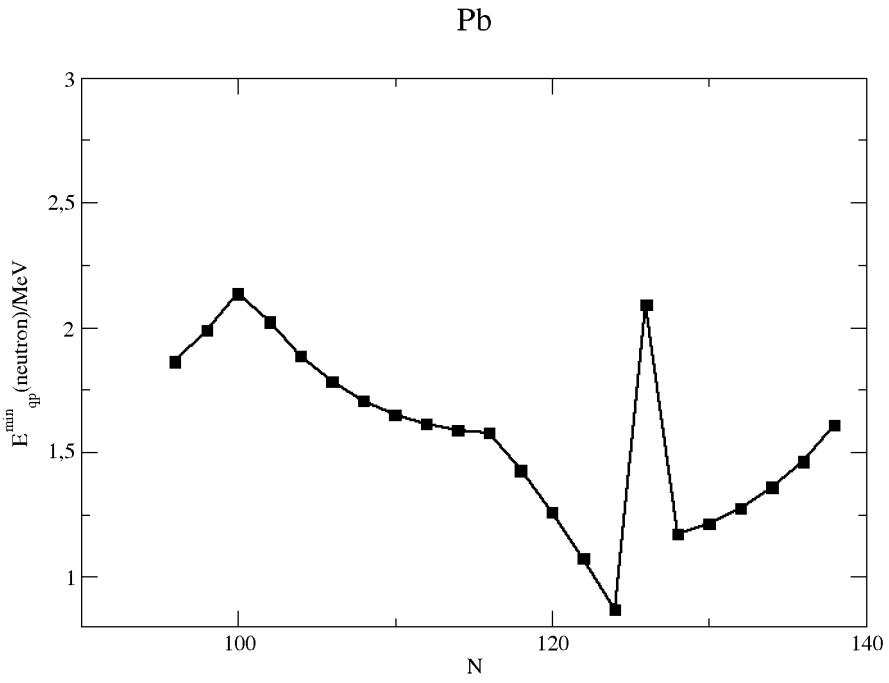


Figure 14: The lowest quasiparticle state at the lowest total BE calculated by HFBTHO for even-even Pb nuclei.

5.3 Calculations on all even-even nuclei

After confirming which parameters need to be tested as the features into the NN, the HFBTHO calculation is performed over all 859 even-even nuclei available from AME16. Figure 15 displays the absolute difference between HFBTHO calculations and AME16 data with different colors labeled in the right bar. The RMS deviation for SLy4 interaction is 6.4597MeV. The HFBTHO has a good performance for nuclei with $Z, N <\sim 90$, but for superheavy nuclei, the deviation can go very high with the magnitude of 10MeV. The poor performance in the superheavy region can be an essential limitation of the HFBTHO code, which makes it unreliable to predict the properties of other heavy nuclei. Therefore, it is necessary to employ the ML means to help improve the results from HFBTHO and extrapolate the algorithm to the whole nuclear chart.

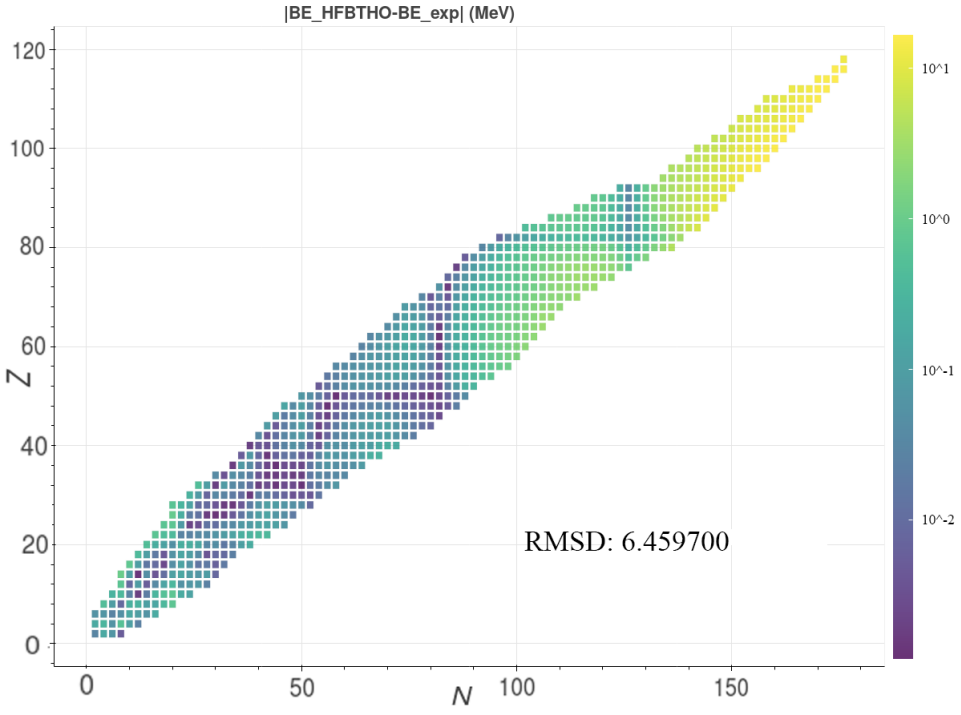


Figure 15: Absolute BE difference between the results for the SLy4 interaction calculated with the HFBTHO program and experimental data.

6 Results from the NN

A test of NN is first applied in the region of even-even Pb isotopes. In one line in the input matrix, proton number Z, neutron number N, a certain deformation β and the corresponding pairing energy , 10 quasiparticle states for neutrons and protons are included to predict the BE corresponded to the β. With 2 hidden layer, 400 hidden neurons and 100 epochs, the training RMS deviation can reach around 4MeV, which is already better than the number from the HFBTHO calculations on all even-even nuclei.

6.1 Predictions on total BE

With the configuration chosen in section 4, the total binding energies from AME16 is first used as the training target. Figure 16 shows how the NN performs with only Z and N as the input features, which means it has no relation to the HFBTHO outputs. The training took 30000 epochs and leads to a training RMS deviation to 3.842949MeV. It improves the HFBTHO results by 40.51%, which means the NN is indeed a possible way to improve the model. However, for some nuclei, the deviations can still reach the magnitude of 10MeV.

However, if one continues to include more features from the HFBTHO codes as the input into the NN, the results tend to be worse. Figure 17(a) shows the results with the lowest HFBTHO total BE added to the features. The RMS deviation reaches 11.922695MeV now. If all the deformations and corresponding BE are added instead of BE_{min}, the RMS deviation turns into an extreme high value around 518MeV. This actually refers to the condition that the NN produces the same value for all outputs, which can also be seen in figure 17(b). The NN remains broken if one tries to add more features.

The reason for the malfunction of the NN is the scaling of the features. The BE used

in the input set is too large compared to the other features. Considering figure 17(a) again, for light nuclei, the BE is smaller and comparable with other feature values, so the overall performance in the light nuclei region is relatively better than that in the heavier region.

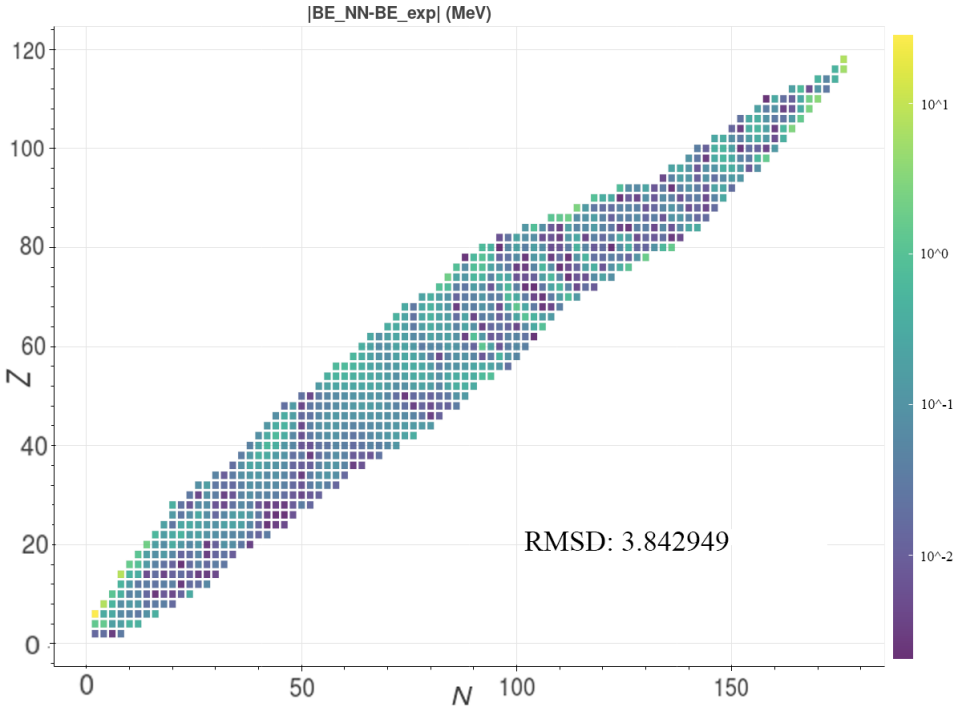
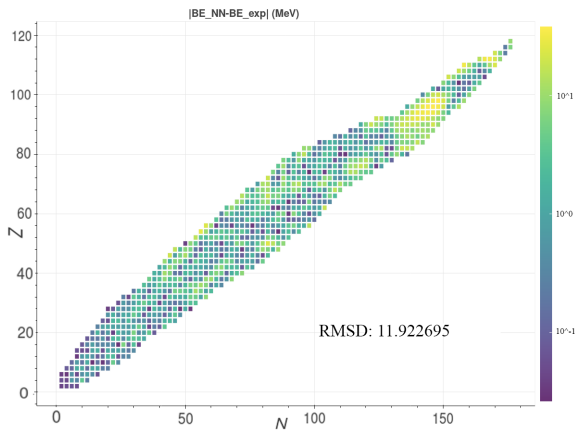
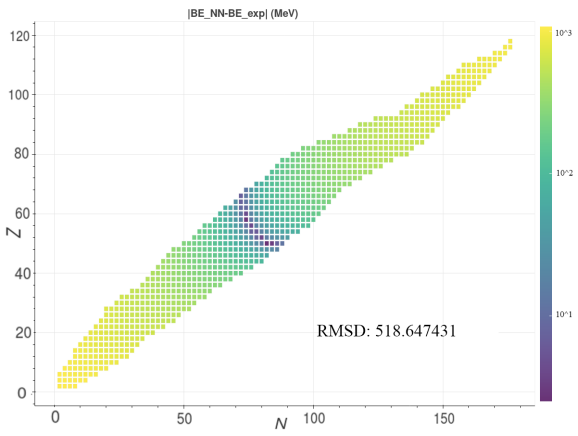


Figure 16: BE difference between the NN and AME16. Training x set: Z, N; training y set: total BE from AME16.



(a) Training x set: Z, N, BE_{min}



(b) Training x set: Z, N, β, $\mathbf{B}E_{\beta}$

Figure 17: BE difference between the NN and AME16 with two different input sets (a) and (b).

6.2 Predictions on BE/A

To solve the problem in section 6.1 , one can simply use the BE per nucleon (BE/A) to substitute the total BE in the NN. Figure 18∼22 show the results with different configurations of the input sets where BE is replaced by BE/A. In these figures, the differences and RMS deviations are still calculated by the theoretical and experimental total BE. In captions, E_{pn} represents the vector of neutron pairing energy, E_{pp} the vector of proton pairing energy, E_{qpn} the vector of neutron quasiparticle states and E_{qpp} the vector of proton quasiparticle states. The lowest RMS deviation 1.422205MeV comes from the input set with Z, N, BE_{min}/A.

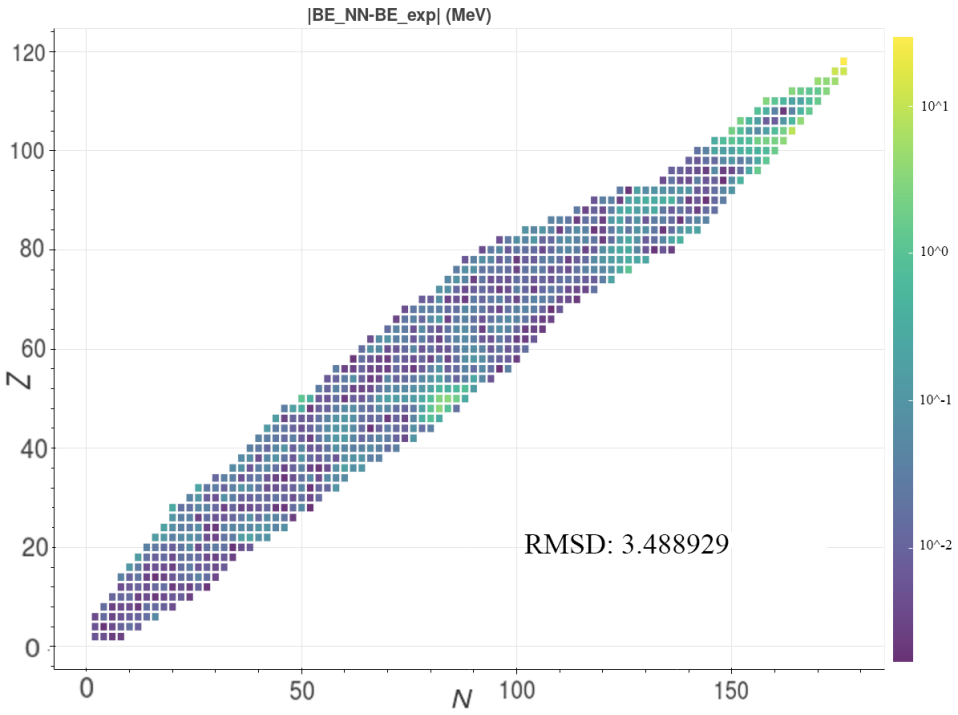


Figure 18: BE difference between the NN and AME16. Training x set: Z, N; training y set: BE/A from AME16.

The BE/A is introduced in the NN to improve the scale of features, as mentioned in section 6.1. This means the appearance of a list of relatively large numbers (from total BE) can be avoided by dividing BE with the number of nucleons. Therefore, the investigations on the NN are based on BE/A in this section.

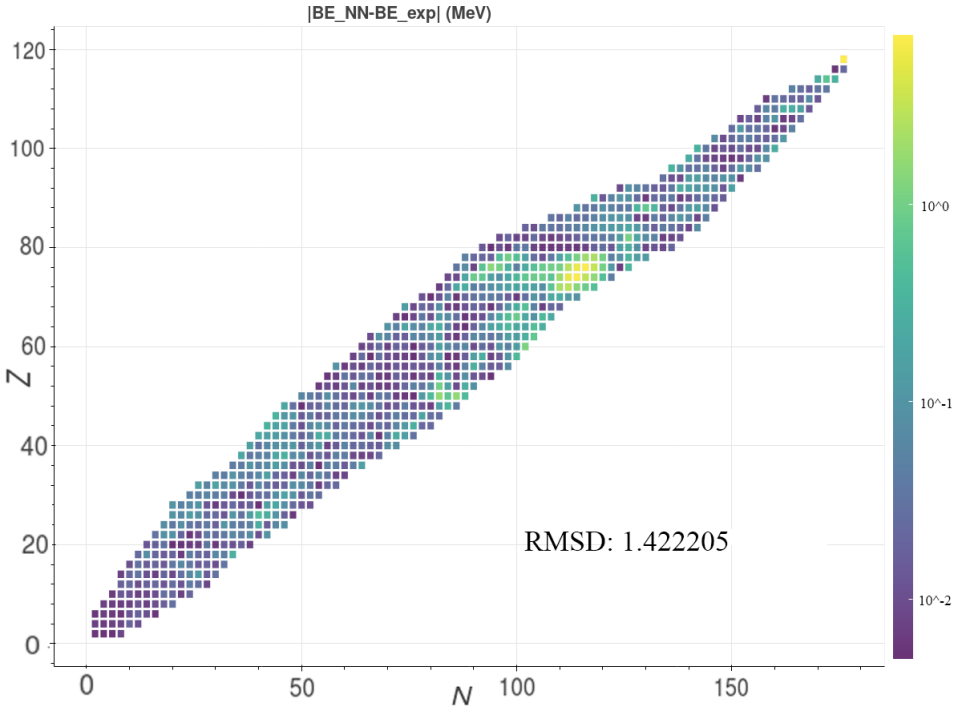


Figure 19: BE difference between the NN and AME16. Training x set: Z, N, BE_{min}/A; training y set: BE/A from AME16.

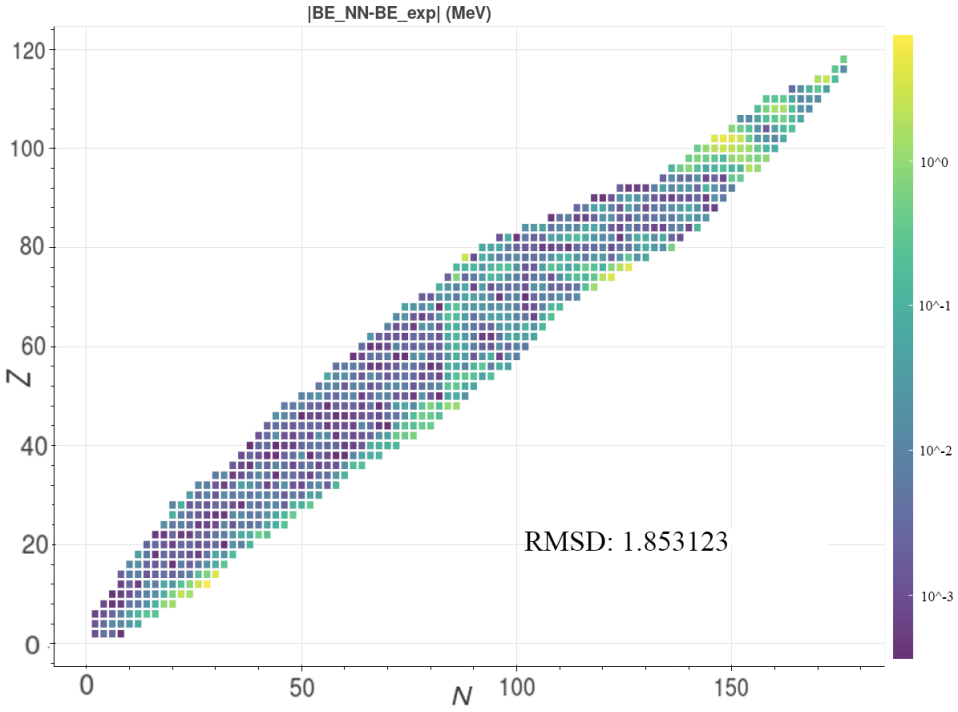


Figure 20: BE difference between the NN and AME16. Training x set: Z, N, β, BE_{β}/A; training y set: BE/A from AME16.

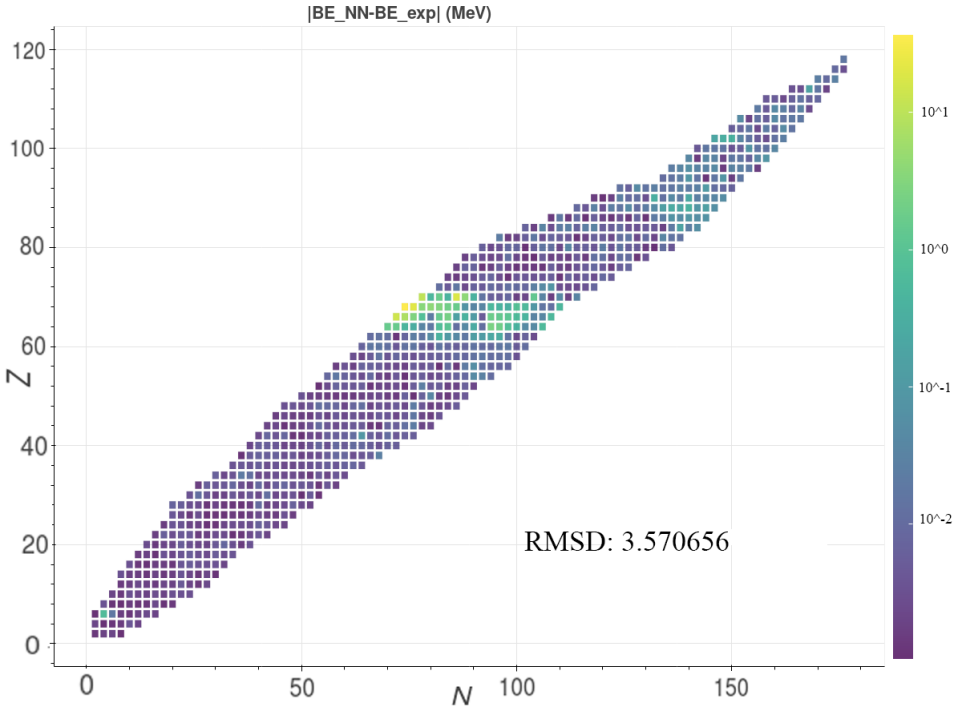


Figure 21: BE difference between the NN and AME16. Training x set: Z, N, β, BE_β/A, E_{pn}, E_{pp}; training y set: BE/A from AME16.

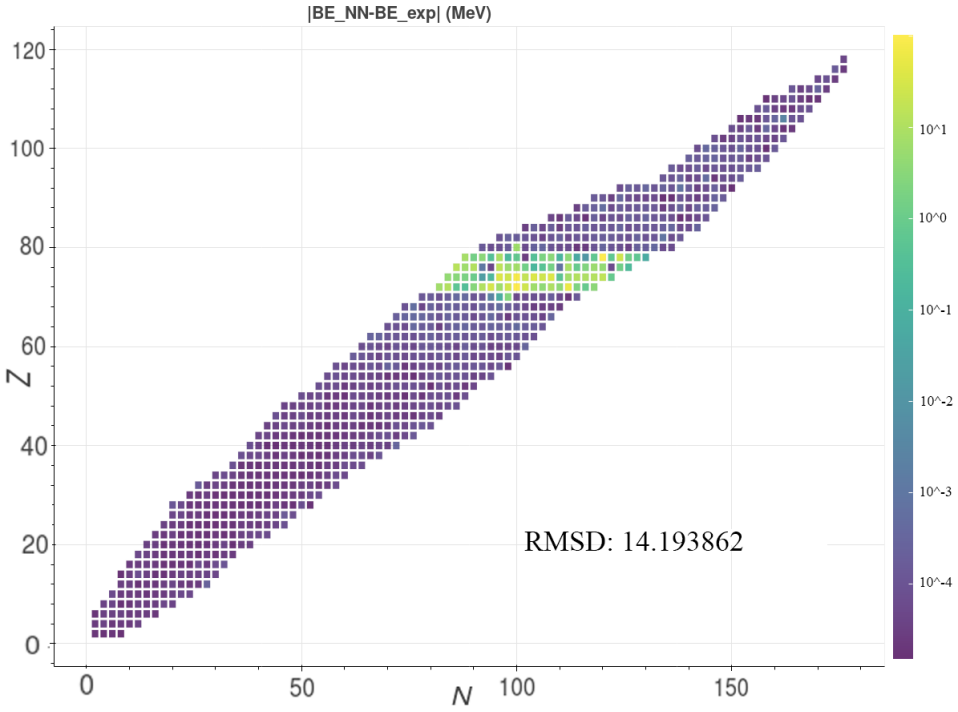


Figure 22: BE difference between the NN and AME16. Training x set: Z, N, β, BE_β/A, E_{pn}, E_{pp}, E_{qpn}, E_{qpp}; training y set: BE/A from AME16.

If one increases the number of features, the RMS deviation will increase rapidly. The training RMS deviation (train_rmsd), validation RMS deviation (rmsd) and total BE RMS deviation (total_rmsd) from the best model in each NN are listed in table 2. For

more features included, the train_rmsd and total_rmsd decrease, while the validation rmsd grows too rapidly. Also from figure 21 and 22, one can observe that the nuclei with large deviations locate closely in one area. This phenomenon results very likely from the cross validation, which means the NN has the overfitting within the training set where the model can fit the training set in a too good way to raise the error in the validation set. Another performance is run on the set from figure 22, where the validation set from the best model is marked with the red frame on the plot, shown in figure 23. This verifies the overfitting guess.

figure	train_rmsd/MeV	rmsd/MeV	total_rmsd/MeV
18	0.025222	0.035019	3.389649
19	0.011053	0.016452	1.105979
20	0.027252	0.012571	1.794534
21	0.008954	0.116615	1.286227
22	0.005812	0.252907	0.662288

Table 2: The training RMS deviation, validation RMS deviation and total BE RMS deviation from the best model during the cross validation process in each NN.

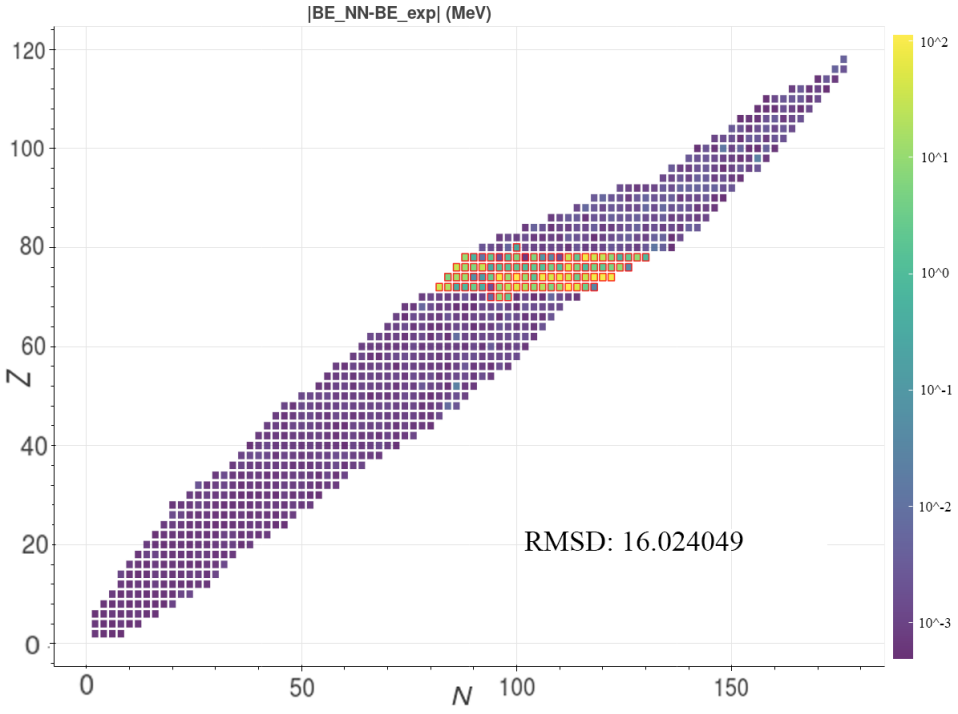


Figure 23: BE difference between the NN and AME16. Training x set: Z, N, β, BE_β/A, E_{pn}, E_{pp}, E_{qpn}, E_{qpp}; training y set: BE/A from AME16. The validation set from the final best model is marked in red square.

One of the ways to improve this kind of overfitting phenomenon is to decrease the number of epochs and to reuse the weights from the last cross validation. If one uses the parameters from last run directly, the NN will produce results shown in figure 24 with only 50 epochs. The RMS deviation can reach 2.866138MeV within very short time, which is a great improvement to the results of 14.193862 MeV in figure 22. The area with large

deviations is somehow extrapolated. In the papers [24, 25], Bayesian neural network is introduced with the similar attempt to reuse the parameters. This optimization deserves more research into it.

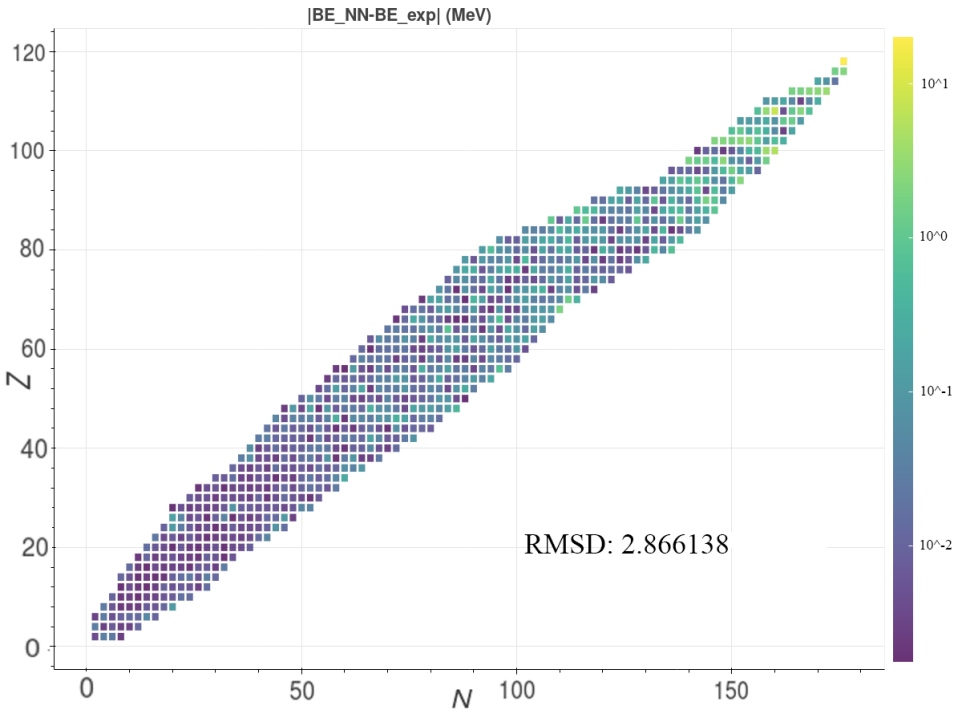


Figure 24: BE difference between the NN and AME16 with 50 epochs for each cross validation. Training x set: Z, N, β, BE_β/A, E_{pn}, E_{pp}, E_{qpn}, E_{qpp}; training y set: BE/A from AME16.

6.3 Predictions on $BE_{HFBTHO} - BE_{exp}$

Among the investigations in the previous sections, the best RMS deviation on the total BE is 1.422205MeV. Compared to the results from [24, 25, 26], it's still under the expectation. One can then use the NN to give predictions on the energy difference $\Delta = BE_{HFBTHO} - BE_{exp}$. For the input set of the NN, apart from Z, N and β, it's also interesting to include the pairing energy because it plays an important role in binding energy, especially when one compares the experimental data to the HFBTHO program which pays much attention to the pairing correlation. The result is shown in figure 25, where no sign of overfitting is found as in figure 22. The RMS deviation can reach 0.218308MeV, which is very small compared to other previous tests. If one add the term Δ_{NN} to BE_{HFBTHO}, the total BE RMS deviation can improve by 96.62% compared the original deviation from HFBTHO. Therefore, this is a potentially good model to predict the BE of all even-even nuclei, but further investigations are needed concerning the applicability of the model to regions with few experimental data such as superheavy regions and neutron drip line.

However, the results from [24, 25, 26] on other theoretical models can bring the RMS deviation below 0.2MeV or even better. Therefore, the NN still needs further modifications and improvements and potential means include involving more features with attempts to avoid overfitting, changing the configurations of the artificial NN or changing to a new ML algorithm entirely.

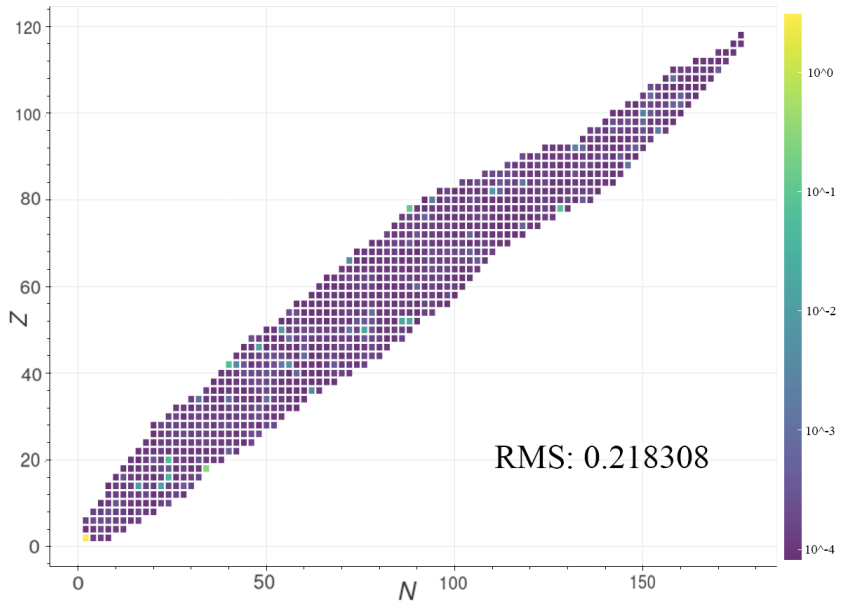


Figure 25: Energy difference between the ΔBE_{NN} and $BE_{HFBTHO} - BE_{exp}$ with 10000 epochs for each cross validation. Training x set: Z, N, β, E_{pn}, E_{pp}, ; training y set: $BE_{HFBTHO} - BE_{exp}$.

Chapter IV
Summary and outlook

The first part of the project is focused on nuclear physics, using the HFBTHO program to calculate all even-even nuclei. First, ^{48}Ca and ^{200}Hg are investigated with different numbers of shells N_{sh} and temperatures T as the input parameters into the program. The best N_{sh} and T are chosen to be 14 and 0.04MeV to achieve the balance between accuracy and speed of the calculations.

The next step is to have an insight to the code performance on all even-even Pb isotopes. 22 nuclei with 21 deformations are calculated by the HFBTHO program with the input parameters determined from the previous investigation. From the plots in figure 11, 12, 13, 14, one can find that the parameters: BE_{HFBYHO} with different deformations, the lowest BE_{HFBTHO}/the difference between the lowest $BE_{HFBTHO}$and experimental BE from AME16, the pairing energies of neutrons and protons and the neutron/proton quasiparticle states can reflect the tendency on how the BE changes with different nuclei. Therefore, these parameters are chosen to be part of the input set into the NN. Finally, the HFBTHO program is run over all 859 even-even nuclei available in AME16. The RMS deviation comes to 6.4597MeV.

With the help of NN, the deviation improves. Different investigations on different selections of the input features into the NN based on Idini's code [3] are summarized through figures 27∼30 in the appendix, where the x-axis is the A number. Compared to figure 26 with $BE_{HFBTHO} - BE_{exp}$, the NN predicting the total BE gives a better RMS deviation of 3.842949MeV. However, this kind of NN is not successful when more features are included. The reason is that the total BE is too large compared to other features. Then, one can train the NN with BE/A. The best model has the input with Z, N and BE_{min} shown in figure 29, where the RMS deviation can reach 1.422205MeV. If one trains the NN on the difference between the lowest BE_{HFBTHO} and BE_{exp}, the best model one can get from this project is with the input of Z,N,β, E_{pn}, E_{pp}, which leads to the RMS deviation of 0.218308MeV and improves the HFBTHO result with 96.62%.

The results show that the ML algorithms can largely improve the HFBTHO calculations. It is quite impressive to lower the RMS deviation from 6.4597MeV to 0.218308MeV. This makes the combination of nuclear physics and ML possible to predict superheavy nuclei and nuclei close to the neutron drip line. Further investigations are required to extrapolate the model to these regions. However, the NN used in this project can still be improved. How to deal with the feature scaling is an interesting topic which can affect the output much. One can also try to solve the overfitting phenomenon with more features involved by tuning the training parameters or changing the type of the NN. Examples include [24, 25, 26].

References

[1] M. Stoitsov, J. Dobaczewski, W. Nazarewicz, and P. Ring, Computer Physics Communications **167**, 43 (2005).

[2] F. Chollet et al., *Keras*, `https://github.com/fchollet/keras` (2015).

[3] A. Idini, arXiv preprint arXiv:1904.00057 (2019).

[4] H. Geiger and E. Marsden, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **82**, 495 (1909).

[5] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Reviews of Modern Physics **75**, 121 (2003).

[6] N. Kaiser, S. Fritsch, and W. Weise, Nuclear Physics A **700**, 343 (2002).

[7] M. Lutz, B. Friman, and C. Appel, Physics Letters B **474**, 7 (2000).

[8] C. F. V. Weizsäcker, Zeitschrift für Physik **96**, 431 (1935).

[9] W. Cui, *Shell-model calculations in ca-region* (2018).

[10] P. Hohenberg and W. Kohn, Physical Review **136**, B864 (1964).

[11] P. Ring and P. Schuck, *The nuclear many-body problem* (Springer Science & Business Media, 2004).

[12] D. R. Hartree, Mathematical Proceedings of the Cambridge Philosophical Society **24**, 89 (1928).

[13] V. Fock, Zeitschrift für Physik **61**, 126 (1930).

[14] J. C. Slater, Physical Review **81**, 385 (1951).

[15] D. Gogny and P.-L. Lions, ESAIM: Mathematical Modelling and Numerical Analysis **20**, 571 (1986).

[16] J. Mattauch, W. Thiele, and A. Wapstra, Nuclear Physics **67**, 1 (1965).

[17] A. Bohr and B. R. Mottelson, *Nuclear structure*, vol. 1 (World Scientific, 1998).

[18] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Physical Review **108**, 1175 (1957).

[19] L. N. Cooper, Physical Review **104**, 1189 (1956).

[20] W. Younes, D. M. Gogny, and J.-F. Berger, in *A Microscopic Theory of Fission Dynamics Based on the Generator Coordinate Method* (Springer, 2019), pp. 3–40.

[21] D. Vautherin and M. Veneroni, Physics Letters B **29**, 203 (1969).

[22] K. Bennaceur and J. Dobaczewski, Computer Physics Communications **168**, 96 (2005).

[23] D. Vautherin, Physical Review C **7**, 296 (1973).

[24] Z. Niu and H. Liang, Physics Letters B **778**, 48 (2018).

[25] L. Neufcourt, Y. Cao, W. Nazarewicz, F. Viens, et al., Physical Review C **98**, 034318 (2018).

[26] H. F. Zhang, L. H. Wang, J. P. Yin, P. H. Chen, and H. F. Zhang, Journal of Physics G: Nuclear and Particle Physics **44**, 045110 (2017).

[27] StackExchange, *What is the difference between a neural network and a deep neural network, and why do the deep ones work better?* (2015), https://stats.stackexchange.com/questions/182734/what-is-the-difference-between-a-neural-network-and-a-deep-neural-network-and-w/184921.

[28] F. Chollet, *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek* (MITP-Verlags GmbH & Co. KG, 2018).

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Journal of Machine Learning Research **12**, 2825 (2011).

[30] B. Gall, P. Bonche, J. Dobaczewski, H. Flocard, and P. H. Heenen, Zeitschrift für Physik A Hadrons and Nuclei **348**, 183 (1994).

[31] J. Terasaki, H. Flocard, P. H. Heenen, and P. Bonche, Nuclear Physics **A621**, 706 (1997).

[32] M. Yamagami, K. Matsuyanagi, and M. Matsuo, Nuclear Physics **A693**, 579 (2001).

[33] M. Stoitsov, W. Nazarewicz, and S. Pittel, Physical Review C **58**, 2092 (1998).

[34] M. Stoitsov, J. Dobaczewski, W. Nazarewicz, S. Pittel, and D. Dean, Physical Review C **68**, 054312 (2003).

[35] M. Stoitsov, N. Schunck, M. Kortelainen, N. Michel, H. Nam, E. Olsen, J. Sarich, and S. Wild, Computer Physics Communications **184**, 1592 (2013).

[36] A. L. Goodman, Nuclear Physics A **352**, 30 (1981).

[37] W. J. Huang, G. Audi, M. Wang, F. G. Kondev, S. Naimi, and X. Xu, Chinese Physics C **41**, 030002 (2017).

[38] M. Wang, G. Audi, F. G. Kondev, W. Huang, S. Naimi, and X. Xu, Chinese Physics C **41**, 030003 (2017).

Appendix

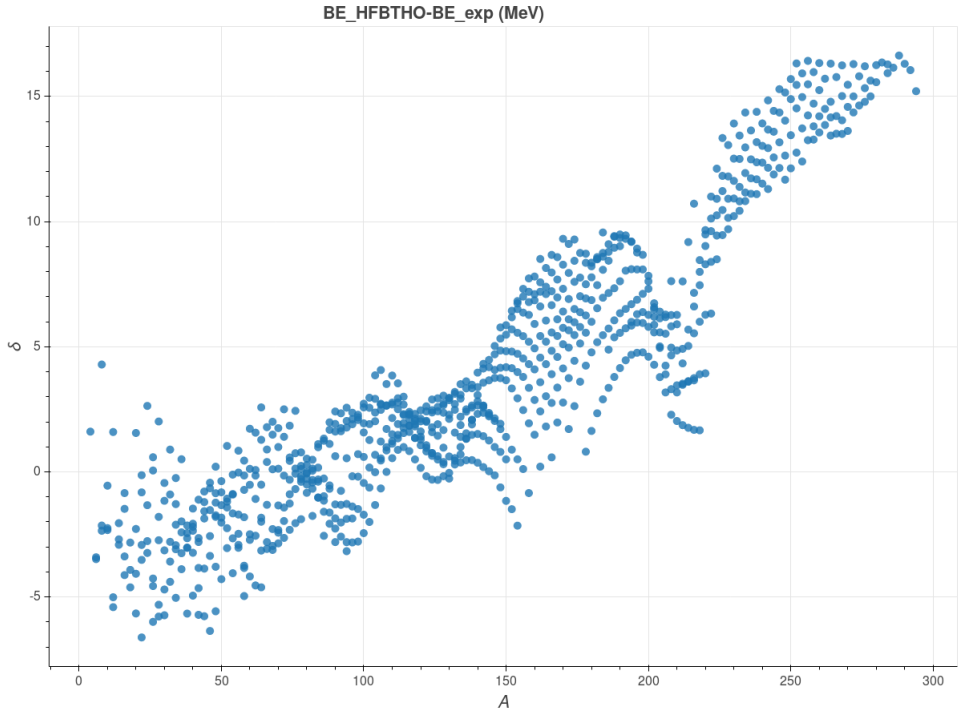


Figure 26: $BE_{HFBTHO} - BE_{exp}$ versus the atomic mass A. This refers to figure 15.

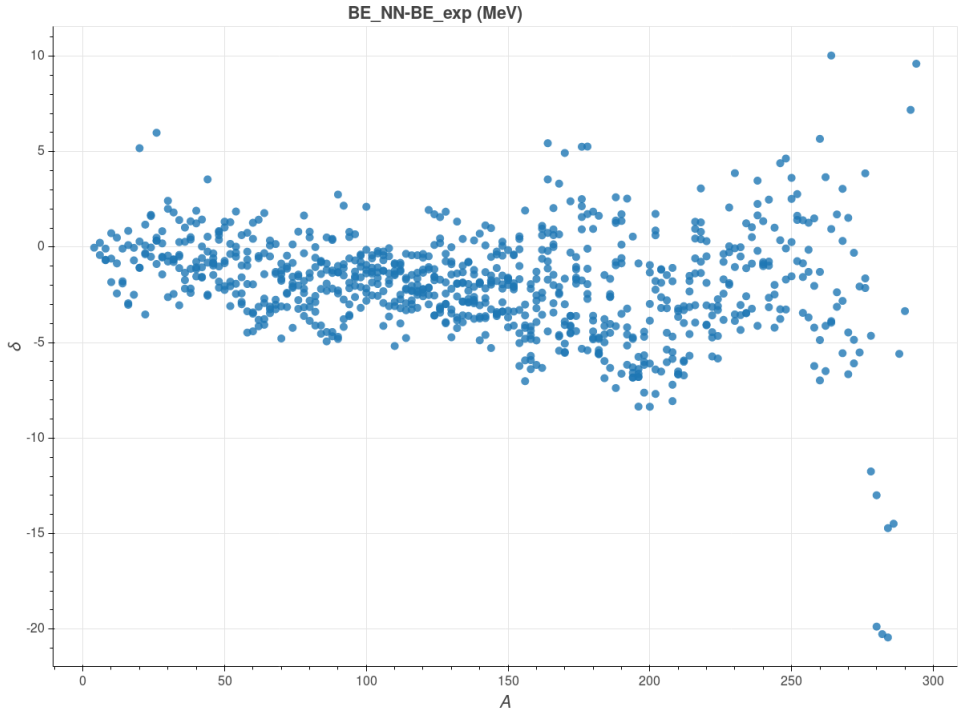


Figure 27: $BE_{NN} - BE_{exp}$ versus the atomic mass A. Training set: x: Z,N; y: BE_{exp}. This refers to figure 16.

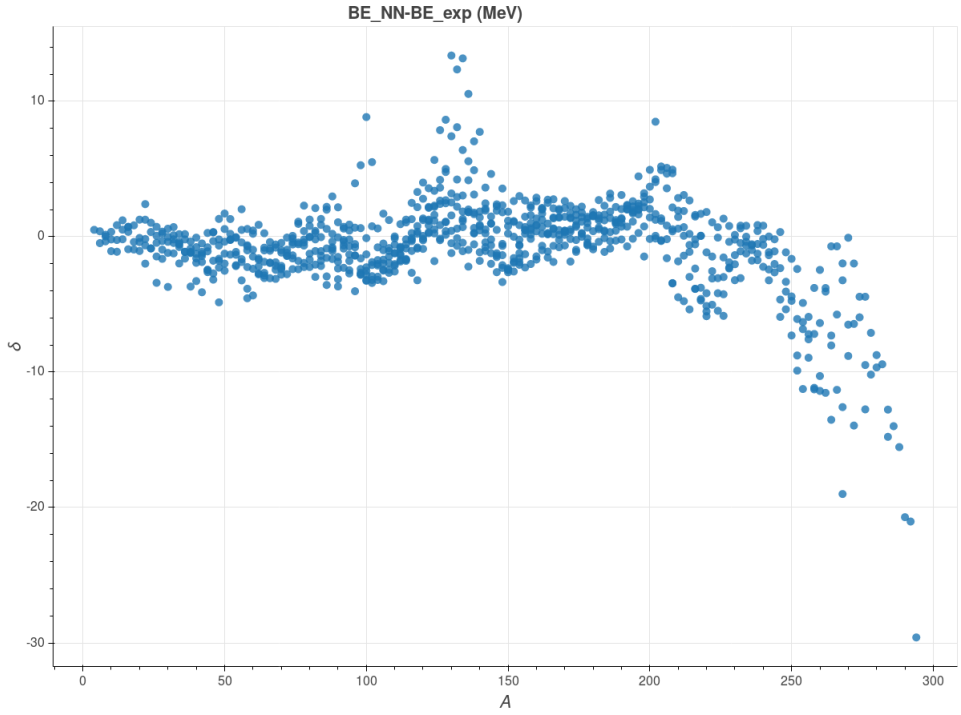


Figure 28: $BE_{NN} - BE_{exp}$ versus the atomic mass A. Training set: x: Z,N; y: BE_{exp}/A. This refers to figure 18.

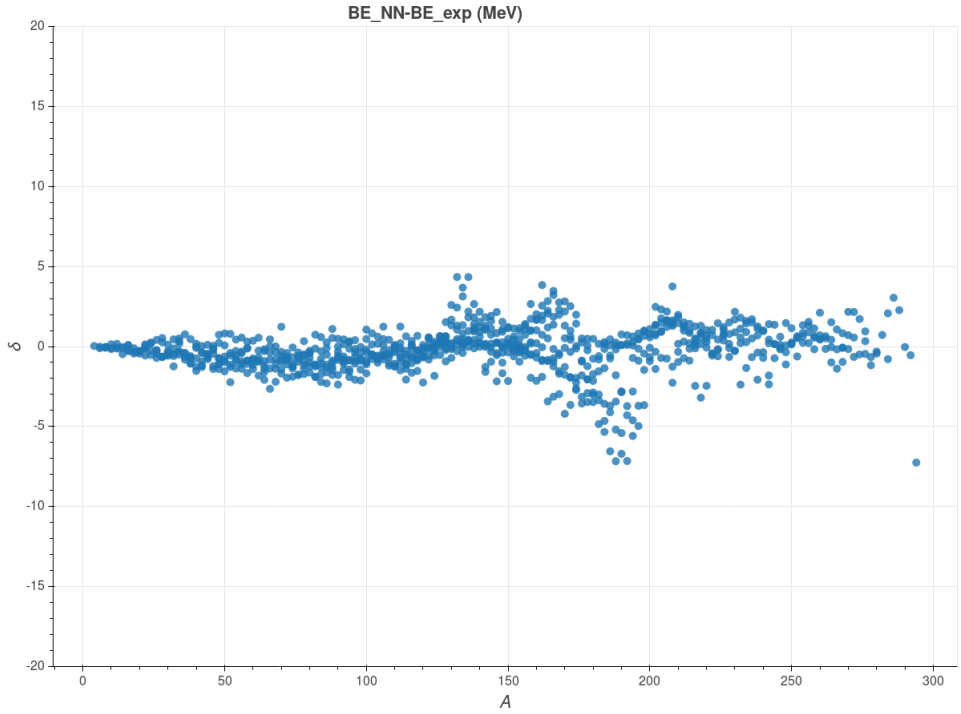


Figure 29: $BE_{NN} - BE_{exp}$ versus the atomic mass A. Training set: x: Z,N,E_{min}; y: BE_{exp}/A. This refers to figure 19.

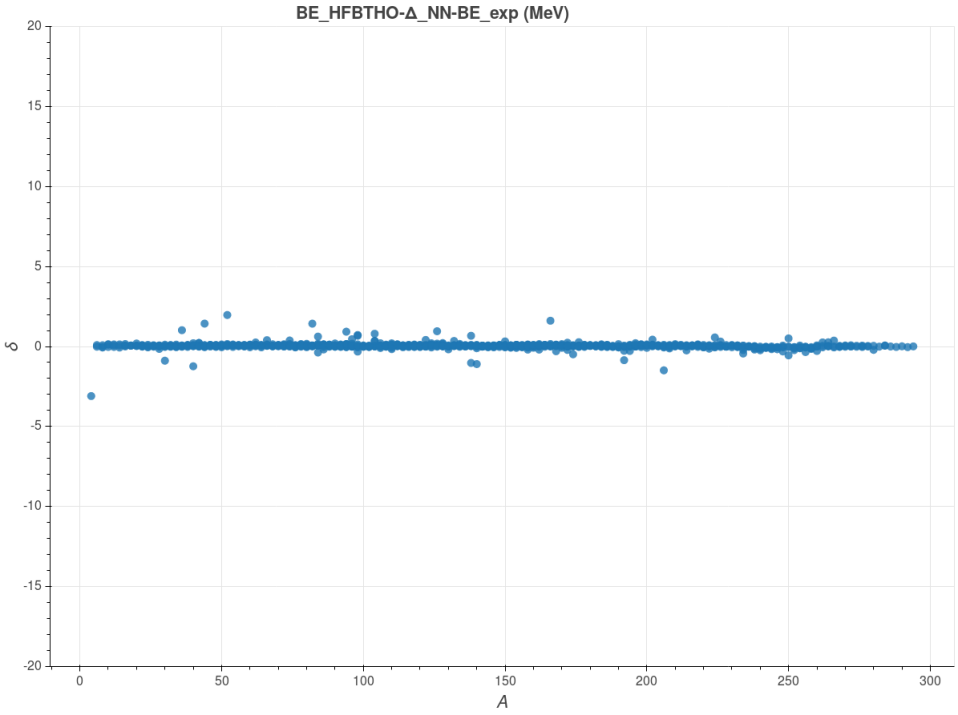


Figure 30: $BE_{HFBTHO} - \Delta_{NN} - BE_{exp}$ versus the atomic mass A. Training set: x: Z,N,β, E_{pn}, E_{pp}; y: $BE_{HFBTHO} - BE_{exp}$. This refers to figure 25.