

Машинное обучение в задаче экстраполяции вариационных расчетов в ядерной физике

А. И. Мазур,¹ Р. Э. Шарыпов,^{1,*} А. М. Широков²

¹Тихоокеанский государственный университет. Россия, 680035, Хабаровск, Тихоокеанская, 136

²Московский государственный университет имени М. В. Ломоносова,
Научно-исследовательский институт ядерной физики имени Д. В. Скобельцына
Россия, 119234, ГСП-1, Москва, Ленинские горы, д. 1, стр. 2

(Поступила в редакцию 16.02.2024; после доработки 07.03.2024; подписана в печать 27.03.2024)

Предложен модифицированный метод машинного обучения, использующий ансамбль искусственных нейронных сетей для экстраполяции на случай бесконечного базиса энергий, полученных в вариационных расчетах, в частности в модели оболочек без инертного кора (МОБИК). Используется новая топология нейронной сети, сформулированы критерии отбора как данных, используемых для обучения, так и обученных нейронных сетей для статистического анализа результатов. Метод апробирован на тестовой задаче экстраполяции энергии основного состояния дейтрона с NN -взаимодействием Nijmegen II и позволяет получать устойчивые результаты с высокой статистической достоверностью. Выполнена экстраполяция расчетов в МОБИК с NN -взаимодействием Daejeon16 энергий основных состояний ядер ${}^6\text{Li}$ и ${}^6\text{He}$.

PACS: 07.05.Mh, 21.10.Dr, 21.60.De УДК: 539.14.

Ключевые слова: ядерная модель оболочек, подходы *ab initio*, машинное обучение, методы экстраполяции вариационных расчетов.

DOI: [10.55959/MSU0579-9392.79.2430202](https://doi.org/10.55959/MSU0579-9392.79.2430202)

ВВЕДЕНИЕ

Модель оболочек без инертного кора (МОБИК) [1] — один из перспективных методов *ab initio* (из первых принципов) описания свойств легких атомных ядер. В МОБИК все нуклоны ядра являются спектроскопически активными и используются только реалистические модели нуклон-нуклонных сил. Реалистические NN -взаимодействия строятся в теории мезонного обмена, но в последнее время все большую популярность приобретают NN -взаимодействия, полученные в рамках киральной эффективной теории поля. Параметры реалистических NN -потенциалов подобраны так, чтобы с высокой точностью воспроизводить имеющиеся данные по нуклон-нуклонному рассеянию и свойствам дейтрона. Детальный обзор разных моделей нуклон-нуклонных сил представлен в работе [2]. В нашей работе мы используем NN -взаимодействие Daejeon16 [3], построенное в рамках киральной эффективной теории поля, в котором свойства вне энергетической поверхности определяются подгонкой к свойствам легких ядер при помощи фазово-эквивалентных преобразований, что позволяет описывать ядра без использования трехнуклонных сил.

Результаты вариационных расчетов использующих осцилляторный базис, включая МОБИК, зависят от двух параметров: от максимального числа учитываемых в расчете осцилляторных кван-

тов возбуждения N_{\max} , которое определяет размер базисного пространства модели, и величины осцилляторной энергии $\hbar\Omega$, являющейся параметром базисных функций. Диагонализация матрицы гамильтониана в модельном пространстве N_{\max} при разных значениях $\hbar\Omega$ позволяет построить зависимости от $\hbar\Omega$ энергии того или иного состояния $E^{var}(N_{\max}, \hbar\Omega)$ для каждого заданного N_{\max} . За результат предсказания энергии в модельном пространстве N_{\max} обычно принимается значение, соответствующее минимуму этой зависимости. В силу вариационного принципа с ростом N_{\max} кривые зависимости энергии от $\hbar\Omega$ все ближе и ближе приближаются сверху к некоторому значению, а сама зависимость от $\hbar\Omega$ становится все более гладкой. В пределе $N_{\max} \rightarrow \infty$ кривая зависимости от $\hbar\Omega$ энергии $E^{var}(N_{\max}, \hbar\Omega)$ вырождается в прямую, положение которой уже не зависит от величины N_{\max} , что и соответствует окончательно сошедшему значению энергии. Поиск этого окончательно сошедшего значения энергии и составляет цель данной работы.

Основная проблема МОБИК — резкое увеличение требуемых вычислительных ресурсов с увеличением N_{\max} , что связано с экспоненциальным ростом количества базисных функций. Современные суперкомпьютеры позволяют проводить расчеты в МОБИК с разумной точностью для легких ядер с массовым числом вплоть до примерно $A \simeq 20$, причем если для легчайших ядер s - и начала p -оболочки (${}^3\text{He}$, ${}^3\text{H}$, ${}^4\text{He}$, ${}^6\text{He}$, ${}^6\text{Li}$) возможны расчеты с $N_{\max} \simeq 20$, то для ядер середины p -оболочки доступны расчеты с $N_{\max} \leq 12-14$.

* E-mail: sharypov@pnu.edu.ru

Для предсказания результатов, соответствующих бесконечному базису, предложены различные методы экстраполяции [4–10], не имеющие, к сожалению, строгого обоснования. Наиболее достоверным представляется метод [11], в котором результаты экстраполяции определяются посредством локализации полюсов S-матрицы для связанных состояний.

Одним из интересных направлений решения подобных задач является использование машинного обучения. На сегодняшний день методы машинного обучения прочно вошли в арсенал как теоретиков, так и экспериментаторов в области ядерной физики (см. обзор [12]). Задачи экстраполяции результатов с помощью методов машинного обучения представляют особый класс задач, до конца не исследованных, и требующий развития новых алгоритмов.

Впервые возможность применения методов машинного обучения к задаче экстраполяции вариационных расчетов на случай бесконечно больших модельных пространств была исследована в работе [13]. Впоследствии было опубликовано еще несколько работ, где экстраполировались результаты расчетов в МОБИК [14–17]. Целью настоящей работы является продолжение исследований в этом направлении. Главные отличия нашего подхода от работы [13] заключаются в том, что мы используем более сложную архитектуру искусственной нейронной сети и более строгие критерии отбора обученных сетей для получения окончательных предсказаний. Как следствие, в нашем подходе нет необходимости делить все используемые данные на обучающий и тестовый наборы: все данные используются для обучения; в нашем подходе отсутствует проблема переобучения сети. Кроме того, что очень важно, отбор обученных сетей на основе сформулированных критериев позволяет существенно повысить статистическую достоверность наших предсказаний.

Основная идея применения машинного обучения к задаче экстраполяции заключается в обучении ансамбля искусственных нейронных сетей на результатах вариационных расчетов для энергии заданного уровня (основного состояния системы в нашем случае) $E^{(var)}(N_{\max}^i, \hbar\Omega^i)$ (индекс i нумерует все используемые сочетания значений N_{\max} , $\hbar\Omega$) в модельных пространствах N_{\max}^i , начиная с некоторого минимального (нижнего) $N_{\max}^i = N_{\max}^l$ вплоть до максимального (верхнего) значения $N_{\max}^i = N_{\max}^u$, и последующей экстраполяции предсказаний на случай достаточно больших модельных пространств $N_{\max} = N_{\max}^f$, где достигается практически полная сходимость расчетов. Входными данными нейронной сети выступают наборы N_{\max}^i и $\hbar\Omega^i$, а выходными — предсказываемые нейронной сетью значения энергии $E^{(pred)}(N_{\max}^i, \hbar\Omega^i)$, которые сопоставляются с полученными в вариационных расчетах соответствующими значениями $E^{(var)}(N_{\max}^i, \hbar\Omega^i)$ в процессе машинного обучения. Обученные нейронные сети подвергаются отбору на основании некоторых

критериев, которые обсуждаются ниже. Каждая из обученных нейронных сетей выдает свои предсказания энергий $E^{(pred)}(N_{\max}^f, \hbar\Omega^i)$, которые в идеале не зависят от $\hbar\Omega^i$. После этого для получения экстраполированного значения энергии и оценки соответствующей погрешности результаты предсказаний прошедших отбор ансамбля нейронных сетей обрабатываются статистическими методами.

Метод экстраполяции, описываемый в настоящей работе, тестируется на модельной вариационной задаче — расчете в осцилляторном базисе энергии основного состояния дейтрона с реалистическим нуклон-нуклонным потенциалом Nijmegen II [18]. Выбор этой задачи в качестве тестовой обусловлен тем, что, во-первых, — это возможность проведения расчетов в очень больших модельных пространствах, так как в этом случае число базисных функций растет линейно с ростом N_{\max} . Во-вторых, известен точный ответ $E_{\text{exp}} = -2.224$ МэВ. В-третьих, отметим медленную сходимость вариационных расчетов (см. рис. 1). В-четвертых, в этой задаче, как видно на рис. 1, отчетливо проявляется так называемый четно-нечетный эффект относительно значений $N_{\max}/2$: результаты расчетов энергий $E^{(var)}(N_{\max}, \hbar\Omega)$ с соседними четными и нечетными значениями $N_{\max}/2$ группируются попарно, что позволяет дополнительно тестировать корректность предсказаний и эффективность метода, применяя его по отдельности к наборам данных с четными или нечетными значениями $N_{\max}/2$, а также к полному набору данных. И, наконец, вариационные расчеты модельной задачи обладают дополнительной интересной особенностью. На кривых зависимости $E^{(var)}(N_{\max}, \hbar\Omega)$ от $\hbar\Omega$ при достаточно малых $N_{\max} < 60$ формируется локальный минимум в области $\hbar\Omega \sim 30\text{--}40$ МэВ, энергии в котором не сходятся к точному значению. При дальнейшем увеличении размеров модельного пространства $N_{\max} > 60$ в области $\hbar\Omega \sim 60$ МэВ начинает формироваться второй минимум, энергии в котором в конечном счете сходятся к точному значению. Таким образом, метод экстраполяции тестируется в случае очень сложного характера сходимости вариационного расчета.

Важную роль в задачах экстраполяции играет предварительный отбор и обработка обучающих данных. Эти вопросы обсуждаются в разд. 1 на примере расчетов энергий основных состояний ядер ${}^2\text{H}$, ${}^6\text{Li}$ и ${}^6\text{He}$.

Процесс обучения нейронной сети характеризуется набором параметров машинного обучения. Выбор топологии нейронной сети и параметров обучения — сложный многоступенчатый процесс, который имеет решающее значение для эффективности метода и который необходимо провести до начала обучения. Эти вопросы рассматриваются в разд. 2. Отметим, что в процессе построения и обучения нейронных сетей в модельной задаче экстраполяции энергии дейтрона мы выбрали топологию и зафиксировали параметры обучения ансамбля нейронных сетей, которые в даль-

нейшем оставались неизменными при исследовании ядер ${}^6\text{Li}$ и ${}^6\text{He}$. Экстраполяцию энергий основных состояний ядер ${}^6\text{Li}$ и ${}^6\text{He}$ мы проводили на основе расчетов в МОБИК с реалистическим NN -взаимодействием Daejeon16 [3].

В третьем разделе рассматривается алгоритм нашего метода экстраполяции, формулируются критерии отбора обученных нейронных сетей и обсуждаются особенности статистической обработки предсказаний прошедших обучение сетей. В четвертом разделе анализируются результаты экстраполяции энергий основных состояний ядер ${}^2\text{H}$, ${}^6\text{Li}$ и ${}^6\text{He}$.

1. ИСХОДНЫЕ ДАННЫЕ: ПРЕДВАРИТЕЛЬНЫЙ ОТБОР И ОБРАБОТКА

На рис. 1, 2, 3 представлены результаты расчетов в МОБИК энергии основного состояния для различных ядер: ${}^2\text{H}$, ${}^6\text{Li}$, и ${}^6\text{He}$.

Для отладки нашего подхода были проведены расчеты энергии основного состояния дейтрона с потенциалом Nijmegen II в модельных пространствах $6 \leq N_{\text{max}} \leq 100$ в интервале значений $\hbar\Omega$ от 10 до 110 МэВ с шагом 1 МэВ. На рис. 1 представлены расчеты в модельных пространствах с $N_{\text{max}} \geq 38$. Как уже отмечалось, в результатах отчетливо проявляется так называемый четно-нечетный эффект: результаты, полученные в соседних модельных пространствах, группируются попарно. Важным для тестовой задачи является и то, что даже для больших значений $N_{\text{max}} = 100$ не достигнута полная сходимость результатов. Мы также уже отмечали, что в этой задаче в зависимости энергии $E^{(var)}(N_{\text{max}}^i, \hbar\Omega)$ от $\hbar\Omega$ формируются два вариационных минимума: один в области $\hbar\Omega \sim 30 - 40$ МэВ при $N_{\text{max}}^i < 60$, однако для $N_{\text{max}}^i \geq 60$ минимум перескакивает в район $\hbar\Omega \sim 60$ МэВ, и именно в области этого минимума вариационная энергия сходится к точному значению.

Расчеты ядер ${}^6\text{Li}$ и ${}^6\text{He}$ в МОБИК требуют существенно больших вычислительных ресурсов. На рис. 2 и 3 представлены результаты для энергий основных состояний этих ядер $E^{(var)}(N_{\text{max}}^i, \hbar\Omega^i)$, полученные с NN -взаимодействием Daejeon16 в модельных пространствах $2 \leq N_{\text{max}}^i \leq 18$ в диапазоне значений $\hbar\Omega^i$ от 8 до 50 МэВ. В больших модельных пространствах ($N_{\text{max}}^i = 18$ в случае ядра ${}^6\text{Li}$ и $N_{\text{max}}^i = 16, 18$ в случае ядра ${}^6\text{He}$) расчеты проведены в более узких диапазонах $\hbar\Omega^i$.

Для обучения нейронных сетей используются не все данные из исходных наборов. В работе [11] было показано, что устойчивые результаты экстраполяции методом SS-HORSE-NCSM достигаются только в случае, если учитывать только данные, которые лежат справа от вариационного минимума зависимости энергии от $\hbar\Omega$ в каждом модельном пространстве. Опыт численных расчетов показал,

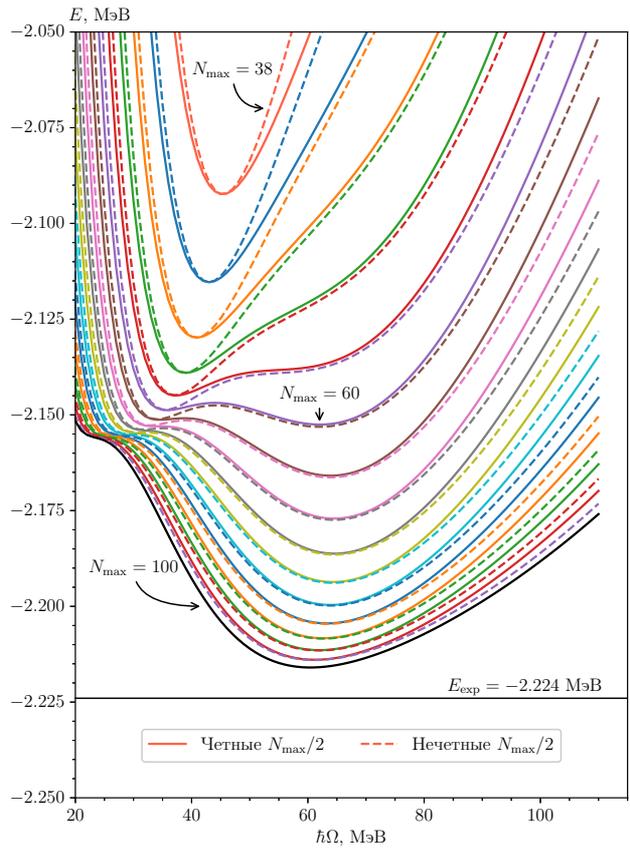


Рис. 1. Результаты расчетов энергии основного состояния ядра ${}^2\text{H}$ в осцилляторном базисе с потенциалом Nijmegen II. Показаны данные начиная с $N_{\text{max}} = 38$. Горизонтальная линия соответствует экспериментальному значению $E_{\text{exp}} = -2.224$ МэВ

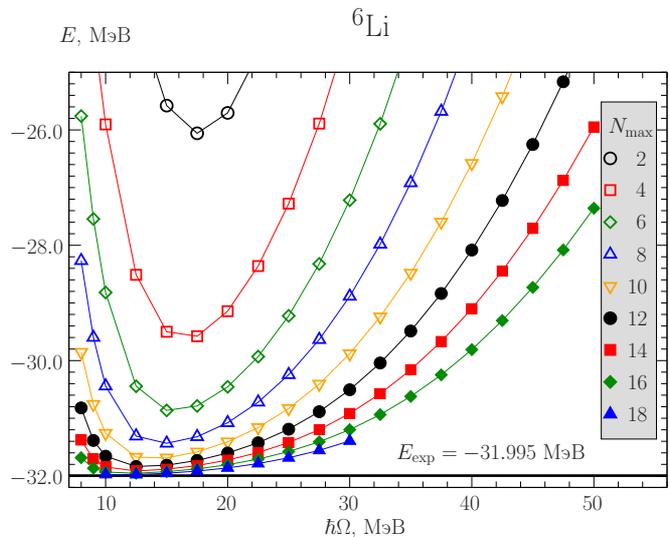


Рис. 2. Результаты расчетов энергий основного состояния ядра ${}^6\text{Li}$ в МОБИК с NN -потенциалом Daejeon16 в модельных пространствах с различными N_{max} . Горизонтальная линия соответствует экспериментальному значению

что подобный отбор данных позволяет существенно улучшить и метод экстраполяции с использо-

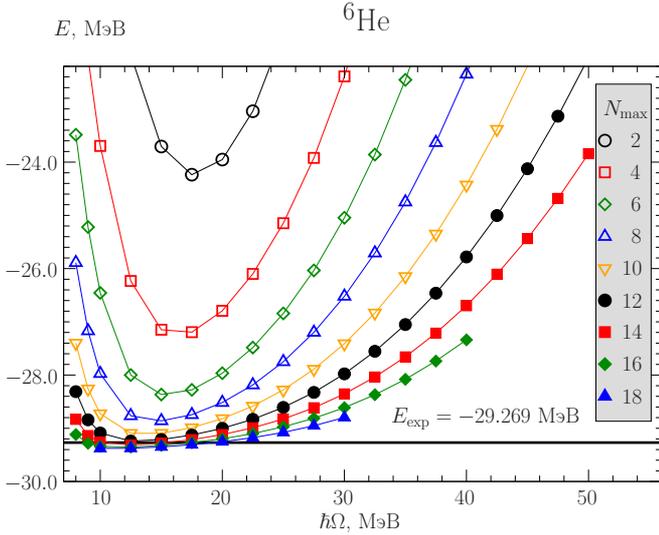


Рис. 3. Результаты расчетов энергий основного состояния ядра ${}^6\text{He}$ в МОБИК с NN -потенциалом Daejeon16 в модельных пространствах с различными N_{\max} . Горизонтальная линия соответствует экспериментальному значению

ванием машинного обучения. Для ядер ${}^6\text{Li}$ и ${}^6\text{He}$ этот критерий отбора используемых данных однозначен. В случае ядра ${}^2\text{H}$ в качестве левой границы отбора по $\hbar\Omega$ мы используем значения, соответствующее локальному минимуму в области $\hbar\Omega \sim 30 - 40$ МэВ для $N_{\max}^i < 60$ и положению второго локального минимума в области $\hbar\Omega \sim 60$ МэВ для $N_{\max}^i \geq 60$, когда энергии вариационных расчетов в нем становятся ниже, чем в первом. Набор входных данных $N_{\max}^i, \hbar\Omega^i$, попадающие в указанные интервалы, и соответствующие им значения вариационных энергий $E^{(var)}(N_{\max}^i, \hbar\Omega^i)$ в дальнейшем будем называть *обучающим набором данных*.

Нашей основной задачей является исследование зависимости предсказаний энергии основного состояния $E^{(pred)}(N_{\max}^f, \hbar\Omega)$ от верхней границы N_{\max}^u значений N_{\max}^i , включенных в обучающий набор данных, при фиксированной нижней N_{\max}^l . В тестовой задаче с дейтроном мы рассматриваем два варианта нижней границы: $N_{\max}^l = 6$ и $N_{\max}^l = 18$. В случае ядер ${}^6\text{He}$ и ${}^6\text{Li}$ нижняя граница $N_{\max}^l = 2$.

Распространенным приемом предобработки обучающего набора данных различной размерности и/или масштаба в машинном обучении является масштабирование. Мы используем простое линейное масштабирование:

$$\tilde{x} = a + \frac{(x - x_{\min})(b - a)}{x_{\max} - x_{\min}}, \quad (1)$$

где \tilde{x} — масштабированное значение той или иной величины из обучающего набора, x — ее исходное значение, а x_{\min} и x_{\max} — соответственно ее минимальное и максимальное значения в обучающем наборе, а a и b — границы масштабированного набора данных. Масштабированные обучающие данные будем обозначать $\tilde{N}_{\max}^i, \tilde{\hbar\Omega}^i$ и $\tilde{E}^{(var)}(\tilde{N}_{\max}^i, \tilde{\hbar\Omega}^i)$. Ниже во всех исследованиях мы полагаем $a = 0$

и $b = 1$, т.е. приводим к одинаковым интервалам $[0, 1]$ все величины $\tilde{N}_{\max}^i, \tilde{\hbar\Omega}^i$ и $\tilde{E}^{(var)}(\tilde{N}_{\max}^i, \tilde{\hbar\Omega}^i)$.

2. ТОПОЛОГИЯ И ПАРАМЕТРЫ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ

Основным элементом искусственной нейронной сети является нейрон, представляющий собой математическую модель, имитирующую поведение биологического нейрона как части нервной системы. Нейрон получает данные извне или от других нейронов и передает обработанный сигнал другим нейронам.

В настоящей работе используется сеть прямого распространения. Такая сеть состоит из нескольких слоев, каждый из которых имеет только односторонние связи со следующим слоем. В ней отсутствуют циклы, то есть сигнал распространяется по слоям строго в направлении от входного слоя к выходному.

В работах [13–15] в задачах экстраполяции применяются нейронные сети с одним скрытым слоем. Мы предлагаем использовать более сложную топологию с тремя скрытыми слоями, в каждом из которых содержится 10 нейронов (см. рис. 4). В работе [19] мы показали, что использование этой (более сложной) топологии позволяет получать более стабильные результаты экстраполяции. В работе [20] в задаче предсказания масс ядер также утверждается, что сеть с тремя скрытыми слоями не только точнее описывает обучающие данные, но и позволяет получать более устойчивые предсказания.

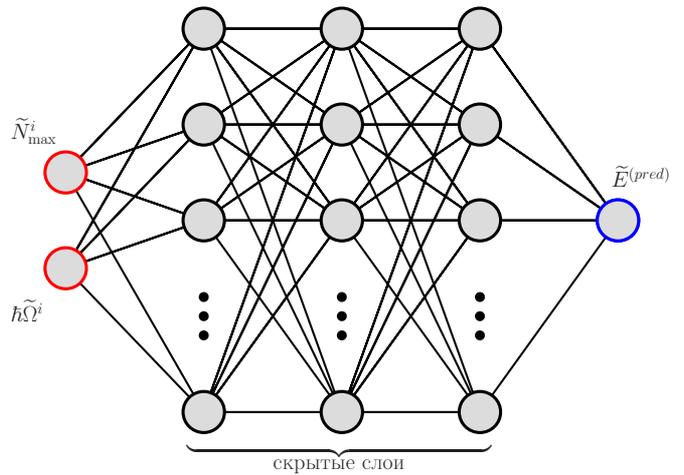


Рис. 4. Топология нейронной сети

Из нейронов входного слоя масштабированные данные \tilde{N}_{\max}^i и $\tilde{\hbar\Omega}^i$ передаются нейронам первого скрытого слоя. Каждый нейрон p этого слоя собирает все входные сигналы x_q и рассчитывает их взвешенную сумму:

$$s_p = \sum_q \omega_{pq} \cdot x_q + b_p. \quad (2)$$

После этого сигнал s_p преобразуется активирующей функцией $f(s_p)$ и передается нейронам следующего слоя, которые, в свою очередь, аналогичным образом обрабатывают поступившие сигналы и передают полученные значения $f(s_{p'})$ далее. Вид активирующей функции в каждом слое может быть различным. Полное прохождение обучающего набора данных для всех значений \tilde{N}_{\max}^i и $\tilde{h}\tilde{\Omega}^i$ по нейронной сети называется эпохой. Важно, что перед обучением каждой сети из ансамбля начальные значения весов ω_{pq} в каждом слое инициализируются случайным образом.

Результаты расчета энергий $\tilde{E}^{(pred)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i)$, полученные на выходном слое, сравниваются с масштабированными энергиями $\tilde{E}^{(var)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i)$ обучающего набора. На основе этого сравнения рассчитывается функция потерь L , что позволяет построить новый набор весов ω_{pq} и сдвигов b_p для каждого слоя.

Прохождение по нейронной сети обучающих данных с новым набором весов и сдвигов неоднократно повторяется. В каждой эпохе ω_{pq} и b_p уточняются таким образом, чтобы минимизировать функцию потерь. Для этого используются оптимизаторы — алгоритмы на основе метода градиентного спуска [21, 22]. Величина изменения весовых коэффициентов и смещений оптимизатором в каждой эпохе пропорциональна скорости обучения η . Обновление набора коэффициентов ω_{pq} и сдвигов b_p происходит каждый раз, когда сквозь нейронную сеть проходит определенное количество данных, которое задается размером пакета B_s (*batch size*). Скорость обучения и размер пакета — важные параметры, влияющие на время и точность обучения нейронной сети.

Процесс обучения нейронной сети при фиксированных параметрах обучения заключается в подборе такого набора весов ω_{pq} и сдвигов b_q — *обучаемых параметров*, которые обеспечивают минимальное отличие выходных результатов $\tilde{E}^{(pred)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i)$ от соответствующих данных $\tilde{E}^{(var)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i)$ обучающего набора. Обученная нейронная сеть с зафиксированными параметрами ω_{pq} и b_p позволяет предсказать энергии $\tilde{E}^{(pred)}(\tilde{N}_{\max}^j, \tilde{h}\tilde{\Omega}^i)$ при значениях $\tilde{N}_{\max}^j > \tilde{N}_{\max}^u$, больших, чем те, на которых нейронная сеть обучалась, то есть экстраполировать энергию связанного состояния в область больших модельных пространств. Энергии с исходной размерностью $E^{(pred)}(N_{\max}^j, h\Omega^i)$ можно получить из масштабированных величин $\tilde{E}^{(pred)}(\tilde{N}_{\max}^j, \tilde{h}\tilde{\Omega}^i)$ с помощью обратного по отношению к (1) преобразования.

Предпринимался поиск оптимальных в смысле величины функции потерь параметров с помощью алгоритма *Tree-Structured Parzen Estimator* [23]. Поиск осуществлялся для одной нейронной сети или с усреднением по небольшому количеству нейронных сетей, т. к. усреднение по всему ансамблю потребовало бы чрезвычайно много времени. Приблизительные диапазоны поиска параметров топологии нейронной сети и параметров машинного обу-

чения составляли:

- число слоев [1, 10];
- число нейронов в слое [4, 30];
- скорость обучения $[10^{-9}, 10^{-4}]$;
- параметр затухания скорости обучения $[10^{-6}, 10^{-4}]$;
- размер обучающего пакета [4, 256];
- число эпох обучения $[10^4, 10^6]$;
- параметры масштабирования данных a и b (см. формулу 1) [0, 10].

Некоторые найденные таким образом параметры были проверены с помощью ансамблирования и использовались для экстраполяции.

Отметим, что скорость обучения и число эпох обучения взаимосвязаны. Как следствие, скорость обучения можно выбрать произвольно в достаточно широком диапазоне значений, а затем соответствующим образом отрегулировать число эпох обучения. Размер обучающего пакета в результате численных экспериментов был выбран равным длине обучающей выборки.

Приведем структуру искусственных нейронных сетей и перечислим конкретные значения параметров машинного обучения, использованные нами.

- Топология нейронных сетей: 2–10–10–10–1, где числа обозначают количество нейронов в слоях, включая входной и выходной (общее число обучаемых параметров ω_{pq} и b_p равно 261).
- Нейроны первого скрытого слоя и выходного слоя обладают линейной активирующей функцией $f(s_p) = s_p$, нейроны второго и третьего — сигмоидальной $f(s_p) = 1/(1 + \exp(-s_p))$.
- Алгоритм, реализующий подстройку весов методом градиентного спуска (оптимизатор), — Adam [24]. Все внутренние параметры оптимизатора, кроме скорости обучения η , установлены по умолчанию.
- Размер обучающего пакета выбран равным числу отобранных обучающих данных $B_s = D$, скорость обучения η — циклическая [25, 26].
- Перед началом обучения веса w_{pq} инициализируются случайно и равномерно распределены в интервале

$$\left[-\frac{\sqrt{6}}{\sqrt{n_p + n_q}}, \frac{\sqrt{6}}{\sqrt{n_p + n_q}} \right],$$

где w_{pq} — вес связи между нейронами соседних слоев p и q , n_p, n_q — число нейронов в этих слоях. Начальные сдвиги b_p полагаются равными нулю.

- Функция потерь — средний квадрат отклонения:

$$L\left(\tilde{E}^{(pred)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i)\right) = \frac{1}{D} \sum_i^D \left(\tilde{E}^{(var)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i) - \tilde{E}^{(pred)}(\tilde{N}_{\max}^i, \tilde{h}\tilde{\Omega}^i) \right)^2. \quad (3)$$

- Число эпох обучения $M_{epoch} = 10^6$.

Отметим, что вблизи выбранных значений параметров ландшафт функции потерь представляется достаточно гладким.

Подбор конкретных значений параметров машинного обучения осуществлялся при рассмотрении тестовой задачи предсказания энергии основного состояния ядра ${}^2\text{H}$. Подчеркнем, что предсказания энергий основных состояний ядер ${}^6\text{Li}$ и ${}^6\text{He}$ были получены с тем же набором параметров машинного обучения.

3. АЛГОРИТМ ЭКСТРАПОЛЯЦИИ И КРИТЕРИИ ОТБОРА

Алгоритм нашего метода экстраполяции энергии связанных состояний и оценка соответствующих погрешностей состоит в следующем.

1. Обучение ансамбля из 1024 нейронных сетей на выбранном наборе обучающих данных. Нейронные сети из ансамбля отличаются лишь начальными значениями весов перед процедурой обучения. Следствием случайной инициализации и стохастического характера процесса обучения является то, что по завершении обучения веса ω_{pq} и смещения b_p принимают разные значения для разных нейронных сетей из ансамбля. Другими словами, локальные минимумы функции потерь, полученные в процессе обучения, оказываются различными для разных сетей.
2. Получение предсказаний энергий $E^{(pred)}(N_{\max}^j, \tilde{h}\tilde{\Omega}^i)$ для каждой нейронной сети из ансамбля для различных значений $N_{\max}^j > N_{\max}^u$ вплоть до некоторого достаточно большого $N_{\max}^j = N_{\max}^f$ для заданного набора значений $\tilde{h}\tilde{\Omega}^i$.
3. Отбор нейронных сетей из ансамбля, предсказания которых удовлетворяют критериям, сформулированным ниже.
4. Статистический анализ той части ансамбля, которая прошла все критерии отбора, для получения результата экстраполяции и оценки погрешности предсказаний.

В наших исследованиях мы не используем, как это зачастую делается, деление входных данных на обучающий, валидационный и тестовый наборы по следующим причинам.

Обучающий набор используется собственно для обучения нейронных сетей, валидационный набор служит для проверки процесса обучения и подбора параметров машинного обучения. В частности, на этом этапе исследуется возможность так называемого переобучения нейронной сети, которое характеризуется достаточно резким увеличением функции потерь на валидационном наборе по сравнению с обучающим набором. Такое поведение функции потерь свидетельствует о том, что нейронная сеть «запоминает» обучающий набор, но при этом теряет способность прогнозировать результаты за пределами обучающего набора. Окончательная проверка работоспособности нейронной сети с уточненными параметрами машинного обучения проводится на тестовом наборе.

На этапе подбора параметров машинного обучения мы исследовали необходимость разбиения данных на обучающий и валидационный наборы. Для выбранной архитектуры нейронной сети было установлено, что в процессе обучения значения функции потерь на валидационной и на обучающей выборке всегда скоррелированы, что свидетельствует об отсутствии проблемы переобучения. Кроме этого, оказалось, что результаты экстраполяций, полученные с разбиением всех имеющихся данных на обучающий и тестовый наборы, практически не отличаются от результатов, полученных с использованием всех данных для обучения. Это подтверждается сходимостью результатов экстраполяций по мере расширения модельного пространства.

Как следствие, в нашем подходе нет необходимости делить все используемые данные на обучающий, валидационный и тестовый наборы: все данные используются для обучения. Использование большого числа данных для обучения фактически приводит к более статистически обоснованным результатам экстраполяции, что особенно важно для будущих расчетов более тяжелых ядер, набор доступных модельных пространств для которых меньше, а сами расчеты в МОБИК требуют больших ресурсов.

Таким образом, в отличие от [13], мы используем весь набор имеющихся данных для обучения нейронных сетей.

Обозначим A^E множество обученных нейронных сетей из ансамбля. Можно сформулировать два физически обоснованных критерия отбора для нейронных сетей, дающих предсказания энергии основного состояния.

Первым физически обоснованным критерием является требование независимости энергии от осцилляторного параметра $\tilde{h}\tilde{\Omega}$ в заданном интервале его значений при больших фиксированных значениях $N_{\max}^i = N_{\max}^f$, когда предполагается, что достигнута сходимость расчетов, то есть

$$E^{(pred)}(N_{\max}^f, \tilde{h}\tilde{\Omega}) = \text{const}. \quad (4)$$

Выбор значения N_{\max}^f зависит от исследуемой задачи и в каждом конкретном случае должен быть протестирован. Так, при экстраполяции энергии основного состояния дейтрона на основе вариационных расчетов в осцилляторном базисе с NN -потенциалом Nijmegen II мы используем значение $N_{\max}^f = 300$, а в случаях ядер ${}^6\text{Li}$ и ${}^6\text{He}$ вполне достаточна величина $N_{\max}^f = 120$.

На практике из-за численных ошибок округления в расчетах проверка равенства (4) должна осуществляться только с некоторой точностью, поэтому мы проверяем неравенство

$$\max \left[E^{(pred)}(N_{\max}^f, \hbar\Omega^i) \right] - \min \left[E^{(pred)}(N_{\max}^f, \hbar\Omega^i) \right] \leq \varepsilon_s, \quad (5)$$

где $\max \left[E^{(pred)}(N_{\max}^f, \hbar\Omega^i) \right]$, $\min \left[E^{(pred)}(N_{\max}^f, \hbar\Omega^i) \right]$ — соответственно максимальное и минимальное значения экстраполированных энергий для всего набора значений $\hbar\Omega^i$. В качестве допустимого отклонения мы выбрали $\varepsilon_s = 0.002$ МэВ. Назовем критерий отбора (5) критерием «прямыны», а множество нейронных сетей, которое ему удовлетворяет, назовем $A_s^E \subset A^E$.

Вторым физическим критерием отбора является вариационный принцип, в соответствии с которым при любых одинаковых значениях $\hbar\Omega^i$ энергия из модельного пространства N_{\max}^i не может лежать выше энергии из модельного пространства $N_{\max}^{i'}$, если $N_{\max}^i > N_{\max}^{i'}$. Вариационный принцип строго соблюдается в расчетах МОБИК, то есть $E^{(var)}(N_{\max}^i, \hbar\Omega^i) \geq E^{(var)}(N_{\max}^{i'}, \hbar\Omega^i)$. Он должен быть справедливым и для результатов экстраполяции в более высокие модельные пространства $N_{\max}^j > N_{\max}^u$.

В нашем подходе критерий отбора по вариационному принципу реализуется следующим образом. Рассмотрим два последовательных модельных пространства с $N_{\max}^{j'} = N - 2$ и $N_{\max}^j = N$. Сравним $E^{(pred)}(N - 2, \hbar\Omega^i)$ и $E^{(pred)}(N, \hbar\Omega^i)$ при каждом значении $\hbar\Omega^i$. Потребуем нестрогого соблюдения неравенства $E^{(pred)}(N_{\max}^{j'}, \hbar\Omega^i) \geq E^{(pred)}(N_{\max}^j, \hbar\Omega^i)$: допустим возможность его небольшого нарушения на величину $\varepsilon_{vp} = 0.1$ кэВ. Обоснованием для введения такого допуска является то, что сами исходные данные имеют точность порядка 1 кэВ и то, что небольшое нарушение вариационного принципа может быть обусловлено численными ошибками в процессе вычислений, особенно при больших N_{\max}^j . Таким образом, проверяется совместное

выполнение неравенств:

$$\begin{cases} E^{(pred)}(N - 2, \hbar\Omega^1) - E^{(pred)}(N, \hbar\Omega^1) \leq \varepsilon_{vp}, \\ E^{(pred)}(N - 2, \hbar\Omega^2) - E^{(pred)}(N, \hbar\Omega^2) \leq \varepsilon_{vp}, \\ \dots \\ E^{(pred)}(N - 2, \hbar\Omega^n) - E^{(pred)}(N, \hbar\Omega^n) \leq \varepsilon_{vp}. \end{cases} \quad (6)$$

Здесь n — число различных значений осцилляторной энергии $\hbar\Omega^i$ в обучающем наборе данных. Если хотя бы одно из неравенств (6) не выполняется для любого $N \leq N_{\max}^f$, то соответствующая нейронная сеть отбрасывается. Полученное с помощью такого отбора множество нейронных сетей назовем A_{vp}^E . В отличие от традиционного вариационного принципа, используемого в [13], будем называть этот критерий «мягким» вариационным принципом.

В качестве еще одного критерия мы использовали отбор по ошибке аппроксимации нейронной сети, под которой понимается величина функции потерь на обучающей выборке по завершении процесса обучения. Из исходного ансамбля исключается 5% нейронных сетей с наибольшей ошибкой аппроксимации. Распределение ошибки аппроксимации, как правило, устроено следующим образом: оно имеет острый «пик» при малых значениях, соответствующий большинству нейронных сетей, и длинный «хвост» при больших значениях, который описывает ошибки аппроксимаций небольшого числа нейронных сетей. Наличие «хвоста» распределения означает, что в процессе обучения был достигнут некий локальный минимум функции потерь, из которого нейронная сеть не смогла выйти. Описанный способ позволяет отбросить нейронные сети с наибольшей ошибкой на обучающей выборке, оставив при этом основную часть ансамбля. Получаемое таким образом множество назовем A_{err}^E . В некоторых работах, например в [14], в процессе решения аналогичной задачи делается упор на мониторинг такой ошибки. Однако наши численные эксперименты показали, что далеко не всегда наблюдается корреляция между величиной ошибки аппроксимации и предсказаниями нейронной сети при $N_{\max} = N_{\max}^f$.

Наконец, в нашем алгоритме используется критерий отбора, основанный на поведении предсказаний при изменении осцилляторного параметра $\hbar\Omega^i$ и на поведении предсказаний при изменении параметра N_{\max}^j . Назовем такой критерий отбора критерием «сходимости». В соответствии с ним сначала находится такое значение $N_{\max}^j \equiv N_{\max}^c$, при котором достигается одновременное выполнение двух условий для данной нейронной сети:

$$\begin{cases} \max \left[E^{(pred)}(N_{\max}^c, \hbar\Omega^i) \right] - \min \left[E^{(pred)}(N_{\max}^c, \hbar\Omega^i) \right] < \varepsilon_1, \\ \min \left[E^{(pred)}(N_{\max}^c, \hbar\Omega^i) \right] - \min \left[E^{(pred)}(N_{\max}^c + 2, \hbar\Omega^i) \right] < \varepsilon_2. \end{cases} \quad (7)$$

Первое условие характеризует независимость энергии основного состояния от осцилляторного параметра $\hbar\Omega^i$ в области асимптотических значе-

ний $N_{\max}^j \equiv N_{\max}^c$ и аналогично условию (5). Второе условие обеспечивает уменьшение разницы между минимумами энергии основного состоя-

ния при использовании последовательных модельных пространств. Затем, исходя из распределения N_{\max}^c , выбираются те нейронные сети, предсказания которых сошлись достаточно быстро при увеличении N_{\max}^j . В условии (7) мы используем $\varepsilon_1 = 0.02$ МэВ и $\varepsilon_2 = 0.001$ МэВ и отбираем 80% наиболее быстро сошедшихся нейронных сетей. Отобранное таким образом множество назовем A_c^E .

Число обученных нейронных сетей, прошедших отбор, определяется пересечением множеств $A_l^E \cap A_{vp}^E \cap A_s^E \cap A_c^E$.

4. РЕЗУЛЬТАТЫ ЭКСТРАПОЛЯЦИИ

Ярко выраженный четно-нечетный эффект в расчетах энергии основного состояния ядра ${}^2\text{H}$ в осцилляторном базисе с реалистическим нуклон-нуклонным взаимодействием Nijmegen II (см. рис. 1) позволяет сравнить результаты экстраполяций, проведенных с разными наборами обучающих данных.

На рис. 5 представлены результаты экстраполяции энергии основного состояния ядра ${}^2\text{H}$, полученные с помощью ансамбля нейронных сетей. В обучающие наборы включены результаты расчетов в осцилляторных базисах, характеризуемых максимальным числом квантов N_{\max}^i от нижней границы N_{\max}^l до верхней границы N_{\max}^u , значения которой отложены по горизонтальной оси. Напомним, что мы используем только те данные, которые находятся в интервале от локального минимума по $\hbar\Omega$ до $\hbar\Omega = 110$ МэВ, как обсуждалось в разд. 1. Рассмотрены случаи, когда в обучающую выборку включены данные с нижней границей $N_{\max}^l = 6$ (пустые символы) и $N_{\max}^l = 18$ (заполненные символы):

1. Верхняя панель — используются данные только с четными значениями $N_{\max}^i/2$.
2. Средняя панель — используются данные только с нечетными $N_{\max}^i/2$.
3. Нижняя панель — используются все данные как с четными, так и нечетными $N_{\max}^i/2$.

Результаты экстраполяций на основе данных из модельных пространств с четными и с нечетными $N_{\max}^i/2$ в целом близки друг к другу. Оба варианта демонстрируют достаточно быструю сходимость. Погрешность предсказаний составляет 3 кэВ в случае четных и 8 кэВ в случае нечетных $N_{\max}^i/2$. Статистическая достоверность результатов в обоих случаях также высокая: число оставшихся после отбора обученных нейронных сетей около 600–700 и не опускается ниже половины исходного числа обучавшихся сетей 1024.

При экстраполяции данных как с только четными, так и с только нечетными $N_{\max}^i/2$, искусственные нейронные сети достаточно хорошо справляются с проблемой двух локальных минимумов

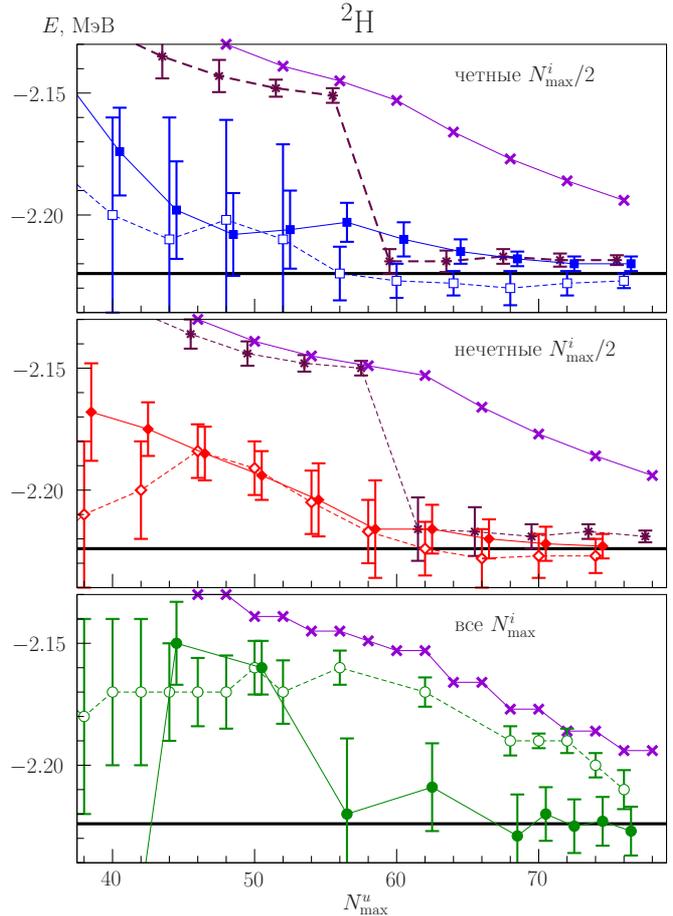


Рис. 5. Экстраполированная энергия основного состояния ядра ${}^2\text{H}$ в зависимости от N_{\max}^u . Верхняя панель — результаты с использованием данных с четными $N_{\max}^i/2$, средняя — с нечетными $N_{\max}^i/2$, нижняя — учитываются все данные. Пустые символы соответствуют расчетам с нижней границей $N_{\max}^l = 6$, закрашенные — с $N_{\max}^l = 18$. Звездочками на двух верхних панелях показаны результаты экстраполяции В [5]. Косые крестики, соединенные линиями, соответствуют вариационным минимумам в модельном пространстве N_{\max}^u . Горизонтальные линии соответствуют точному значению энергии основного состояния дейтрона $E = -2.224$ МэВ

в семействе кривых на рис. 1. Сокращение числа обучающих данных за счет повышения значения N_{\max}^l с 6 до 18, что уменьшает число данных, «сходящихся» к первому локальному минимуму в районе $\hbar\Omega \sim 30\text{--}40$ МэВ, практически не влияет на результаты с $N_{\max}^u > 54$; при меньших значениях N_{\max}^u увеличивается погрешность предсказаний с $N_{\max}^l = 6$, но при этом их диапазоны перекрываются.

Использование всех данных (нижняя панель рис. 5) также приводит к разумным результатам экстраполяции. Однако следует отметить более медленную сходимость и заметно большую погрешность предсказаний. Подчеркнем высокую статистическую достоверность предсказаний и в этом случае: число нейронных сетей после отсева составляет ~ 600 .

Звездочками на двух верхних панелях рис. 5 для сравнения представлены результаты феноменологической экстраполяции В [5], полученные на основе данных с соответственно только четными или только нечетными значениями $N_{\max}/2$ (в силу четно-нечетного эффекта экстраполяция В неприменима для всех данных). Результаты экстраполяций В на основе данных с $N_{\max}^u < 60$ сходятся к некоторому значению $E \sim -2.15$ МэВ вблизи первого локального минимума (см. рис. 1), далекому от точного значения. При учете данных с $N_{\max}^u \geq 60$ результаты экстраполяций В как в случае четных значений $N_{\max}^u/2$, так и в случае нечетных $N_{\max}^u/2$, скачком приближаются к точному значению, но в дальнейшем, по мере роста N_{\max} , не демонстрируют сходимости к точному значению в пределах погрешности. Таким образом, во всех рассмотренных случаях предлагаемый метод на основе машинного обучения нейронных сетей дает более точные предсказания по сравнению с феноменологической экстраполяцией В.

Ядро ${}^6\text{Li}$ для нас представляет особый интерес. Дело в том, что в работе [13] впервые был предложен и применен к этому ядру метод экстраполяции энергии на основе машинного обучения. В качестве используемых данных в [13], как и у нас, используются результаты расчетов в МОБИК с NN -взаимодействием Daejeon16, но алгоритмы экстраполяции различаются. В частности, в [13] топология нейронной сети более простая и содержит только один скрытый слой. Далее, в процессе обучения нейронной сети все используемые результаты МОБИК в [13] разбиваются на два набора — обучающий и тестовый; мы же в процессе обучения используем все отобранные результаты для данного N_{\max}^u .

В обоих подходах происходит обучение ансамбля сетей. Но в нашем случае сформулированы четкие критерии отбора обученных сетей, число отобранных из исходных 1024 сетей в каждом из расчетов с различными N_{\max}^u разное, но, как видно из табл. 1, не меньше 800. Исключение составляют только расчеты с $N_{\max}^u = 10$ — в этом случае число отобранных сетей составляет 443, что, по всей видимости, обусловлено слишком маленьким числом обучающих данных. В подходе [13] стартовый ансамбль содержит 400 нейронных сетей. В процедуре отбора это число может увеличиться. Из этого ансамбля отбирается только 50 лучших нейронных сетей — значительно меньше, чем в нашем подходе. Как следствие, статистическая достоверность предсказаний и оценок погрешностей в нашем подходе выше.

Значения энергий, отвечающих минимумам вариационных расчетов в разных модельных пространствах МОБИК, результаты экстраполяции В [5], а также предсказания, полученные в подходах [13] и нашем, приведены в табл. 1 и представлены на рис. 6.

Отметим, что феноменологическая экстраполяция В не имеет в своей основе никаких предполо-

жений о структуре волновой функции, но, несмотря на свою простоту, для энергий работает не хуже, а зачастую лучше других, более сложных феноменологических экстраполяций. Кроме того, в работе [13] проводится сравнение именно с экстраполяцией В, поэтому для нас результаты этой экстраполяции являются определенным репером.

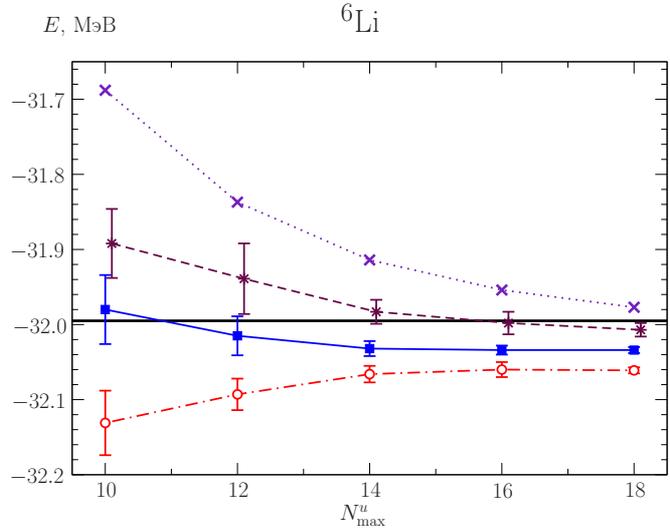


Рис. 6. Экстраполированная энергия основного состояния ядра ${}^6\text{Li}$ на основе расчетов в МОБИК с NN -взаимодействием Daejeon16. Косыми крестиками показаны значения минимума вариационных расчетов в соответствующем модельном пространстве N_{\max}^u , звездочки соответствуют результатам экстраполяции В [5], незакрашенные кружки — результатам работы [13], закрашенные квадратики — нашим результатам. Вертикальными линиями показаны соответствующие погрешности. Горизонтальная линия соответствует экспериментальному значению $E_{\text{exp}} = -31.995$ МэВ

Наши предсказания демонстрируют хорошую сходимость по мере роста N_{\max}^u . Погрешность результатов небольшая, по сравнению с точностью расчетов в МОБИК, составляющих 1 кэВ, и не превышает соответствующих значений в подходе [13]. Наши предсказания лежат ниже результатов, полученных экстраполяцией В, но выше предсказаний работы [13], причем интервалы погрешностей всех трех методов в расчетах с $N_{\max}^u \geq 12$ не перекрываются. Иными словами, полученные тремя способами результаты экстраполяций различимы. Таким образом, два подхода, основанные на методах обучения нейронных сетей — наш и представленный в работе [13], приводят к разным предсказаниям. Но статистическая достоверность наших результатов выше.

В табл. 2 и на рис. 7 представлены результаты экстраполяции энергии основного состояния ядра ${}^6\text{He}$, которое до настоящего времени не исследовалось в подходах с применением нейронных сетей. Как видно, предсказания энергии основного состояния, полученные в нашем подходе, практически не меняются с ростом N_{\max}^u , а погрешность предсказаний уменьшается, хотя остается заметно больше

Таблица 1. Энергии, отвечающие минимумам вариационных расчетов в разных модельных пространствах МОБИК, результаты экстраполяции В, предсказания, полученные в подходе [13] и нашем подходе для ядра ${}^6\text{Li}$. В круглых скобках приведены оценки соответствующих погрешностей; в последней колонке — число отобранных в процессе обучения нейронных сетей

| N_{\max}^u | Вар. минимум, E | Экстрап. В, $E(\Delta E)$ | Рез-ты [13], $E(\Delta E)$ | Наши расчеты, $E(\Delta E)$ | Число отобр. сетей |
|--------------|----------------------|------------------------------|-------------------------------|--------------------------------|-----------------------|
| 10 | -31.688 | -31.892(46) | -32.131(43) | -31.980(46) | 443 |
| 12 | -31.837 | -31.939(47) | -32.093(21) | -32.015(26) | 810 |
| 14 | -31.914 | -31.983(16) | -32.066(11) | -32.032(10) | 830 |
| 16 | -31.954 | -31.998(15) | -32.060(10) | -32.034(6) | 860 |
| 18 | -31.977 | -32.007(9) | -32.061(4) | -32.034(4) | 877 |

Таблица 2. Энергии, отвечающие минимумам вариационных расчетов в разных модельных пространствах МОБИК, результаты экстраполяции В и предсказания, полученные в нашем подходе для ядра ${}^6\text{He}$. В круглых скобках приведены оценки соответствующих погрешностей; в последней колонке — число отобранных в процессе обучения нейронных сетей

| N_{\max}^u | Вар. минимум, E | Экстрап. В, $E(\Delta E)$ | Наши расчеты, $E(\Delta E)$ | Число отобр. сетей |
|--------------|----------------------|------------------------------|--------------------------------|-----------------------|
| 10 | -29.095 | -29.31(11) | -29.48(12) | 284 |
| 12 | -29.241 | -29.36(9) | -29.46(6) | 382 |
| 14 | -29.314 | -29.39(3) | -29.46(5) | 422 |
| 16 | -29.353 | -29.40(2) | -29.45(3) | 417 |
| 18 | -29.377 | -29.41(1) | -29.45(2) | 395 |

(20 кэВ), чем в случае ядра ${}^6\text{Li}$ (4 кэВ). Количество нейронных сетей, остающихся после отбора, также заметно меньше и составляет от 280 до 420 для различных N_{\max}^u . Видимо, снижение точности и статистической достоверности предсказаний обусловлено более медленной сходимостью результатов расчетов в МОБИК.

Как видно из табл. 1 и 2, результаты экстраполяции энергий основных состояний ядер ${}^6\text{Li}$ и ${}^6\text{He}$ стабильны уже при $N_{\max}^u = 10, 12$; дальнейшее увеличение обучающей выборки приводит только к уменьшению неопределенности предсказаний. Таким образом, можно надеяться, что предлагаемый подход будет работать и для более тяжелых ядер p -оболочки. Дополнительным аргументом является то, что на результаты обучения, как оказалось, практически не влияют ни «прореживание» данных, ни увеличение их числа с помощью интерполяции данных для получения энергий при дополнительных значениях $\hbar\Omega$ при данном N_{\max}^u .

ЗАКЛЮЧЕНИЕ

Модифицирован метод экстраполяции на случай бесконечного базиса с помощью машинного обучения для энергии основного состояния на основе обучающих данных из расчетов в МОБИК, предложенный в работе [13]. Основная идея предлагаемой модификации метода заключается в обучении ансамбля искусственных нейронных сетей на основе предварительного отбора обучающих данных, последующем отборе обученных нейронных сетей в соот-

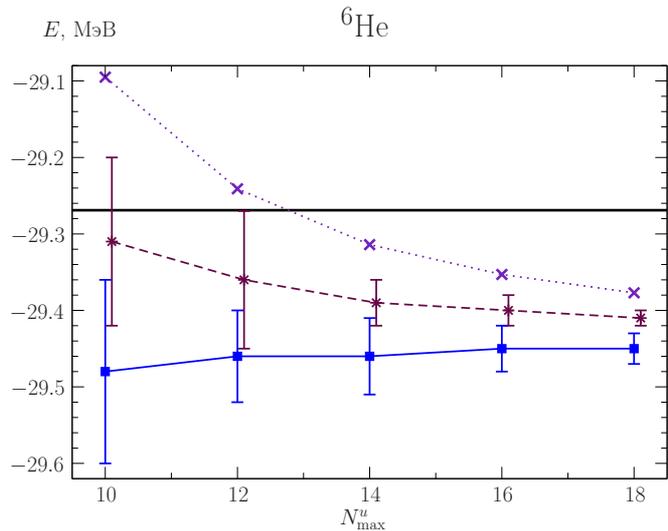


Рис. 7. Экстраполированная энергия основного состояния ядра ${}^6\text{He}$ на основе расчетов в МОБИК с NN -взаимодействием Daejeon16. Обозначения такие же, как на рис. 6. Экспериментальное значение энергии основного состояния $E_{\text{exp}} = -29.269$ МэВ

ветствии с определенными критериями и статистической обработке предсказаний сетей, прошедших отбор. Алгоритм экстраполяции реализован на языке программирования Python с использованием открытых библиотек машинного обучения (Keras [27], TensorFlow [28], TensorFlow Addons [29]).

Предложена новая топология нейронной сети, определены параметры обучения. Модифицирован-

ный метод позволяет получать устойчивые результаты, не требует разбиения используемых данных на обучающий и тестовый наборы, обеспечивает высокую сходимость предсказаний по мере включения в обучающую выборку данных из более высоких модельных пространств, а также высокую статистическую достоверность окончательных результатов.

Для получения предсказаний в случае бесконечного базиса МОБИК и оценки их погрешностей сформулированы физически обоснованные критерии отбора обученных сетей из первоначального ансамбля: «мягкий» вариационный принцип и критерий «прямызны», то есть независимости предсказаний от $\hbar\Omega$ при больших значениях N_{\max}^j . Также сформулированы критерии отбора, основанные на внутренних свойствах ансамбля нейронных сетей: критерий, основанный на величине функции потерь на обучающей выборке по окончании процесса обучения, и критерий, основанный на особенностях сходимости предсказаний при последовательном увеличении N_{\max}^j .

Алгоритм апробирован на тестовой задаче экстраполяции энергии основного состояния дейтрона в вариационных расчетах в осцилляторном базисе с реалистическим NN -взаимодействием Nijmegen II, имеющей точное решение. Проведен анализ влияния отбора обучающих данных на результаты предсказаний. Показано, что для получения устойчивых результатов экстраполяции энергии основного состояния в качестве обучающих данных необходимо использовать только результаты, полученные в вариационных расчетах с $\hbar\Omega^i$, которые находятся правее локального минимума на кривых зависимости $E^{(var)}(N_{\max}^i, \hbar\Omega)$ от $\hbar\Omega$ в каждом модельном пространстве N_{\max}^i .

Результаты вариационных расчетов энергии основного состояния дейтрона с нуклон-нуклонным взаимодействием Nijmegen II характеризуются сильным четно-нечетным эффектом (результаты модельных пространств с четными и нечетными соседними значениями $N_{\max}/2$ группируются попарно) и наличием двух локальных минимумов, формируемых кривыми $E^{(var)}(N_{\max}^i, \hbar\Omega^i)$ в разных

модельных пространствах, причем в одном из них энергии не сходятся к точному значению. Предсказания, полученные на основе вариационных результатов только из «четных» или только из «нечетных» модельных пространств, демонстрируют хорошую сходимость относительно роста N_{\max}^u и практически совпадают друг с другом. Использование для машинного обучения всех результатов из «четных» и «нечетных» модельных пространств также приводит к правильному результату, однако погрешность предсказаний в этом случае больше.

Выполнена экстраполяция расчетов в МОБИК с NN -взаимодействием Daejeon16 энергии основного состояния ядра ${}^6\text{Li}$. Результаты предсказаний демонстрируют хорошую сходимость по мере роста N_{\max}^u . Наши предсказания находятся ниже результатов феноменологической экспоненциальной экстраполяции В, но выше предсказаний, полученных в работе [13] на основе другого алгоритма машинного обучения. При одинаковой точности предсказаний наш подход обладает более высокой степенью статистической достоверности по сравнению с [13].

С помощью разработанного алгоритма на основе расчетов МОБИК с NN -взаимодействием Daejeon16 проведена также экстраполяция энергии основного состояния ядра ${}^6\text{He}$. Исследована сходимость предсказаний по мере включения в обучающую выборку данных из более высоких модельных пространств. Точность предсказаний и статистическая достоверность полученных предсказаний несколько хуже по сравнению с результатами для ядра ${}^6\text{Li}$, что, по всей видимости, обусловлено более медленной сходимостью расчетов МОБИК.

Мы полагаем, что предлагаемый метод экстраполяции является довольно универсальным, и его можно применить к другим наблюдаемым, таким как квадрупольный момент, вероятности электромагнитных переходов в ядрах и др.

Работа выполнена при поддержке гранта РФФИ № 24-22-00276.

-
- [1] Barret R., Navrátil P., Vary J.P. // *Prog. Part. Nucl. Phys.* **69**. 131. (2013).
- [2] Machleidt R. // *Int. J. Mod. Phys. E.* **26**, N 11. 1730005. (2017).
- [3] Shirokov A.M., Shin I.J., Kim Y. et al. // *Phys. Lett. B.* **761**. 87 (2016).
- [4] Zhan H., Nogga A., Barrett B.R. et al. // *Phys. Rev. C.* **69**. 034302 (2004).
- [5] Maris P., Vary J.P., Shirokov A.M. // *Phys. Rev. C.* **79**. 014308 (2009).
- [6] Coon S.A., Avetian M.I., Kruse M.K.G. et al. // *Phys. Rev. C.* **86**. 054002 (2012).
- [7] Maris P., Vary J.P. // *Int. J. Mod. Phys. E.* **22**. 1330016 (2013).
- [8] Kruse M.K.G., Jurgenson E.D., Navratil P. et al. // *Phys. Rev. C.* **87**. 044301 (2013).
- [9] More S.N., Ekstrom A., Furnstahl R.J. et al. // *Phys. Rev. C.* **87**. 044326 (2013).
- [10] Furnstahl R.J., More S.N., Papenbrock T. // *Phys. Rev. C.* **89**. 044301 (2014).
- [11] Широков А.М., Мазур А.И., Куликов В.А. // *ЯФ.* **84**, N 2. 111 (2021).
- [12] Boehnlein A., Diefenthaler M., Fanelli C. et al. // *Rev. Mod. Phys.* **94**. 031003 (2022).
- [13] Negoita G.A., Vary J.P., Luecke G.R. et al. // *Phys. Rev. C.* **99**. 054308 (2019).
- [14] Jiang W.G., Hagen G., Papenbrock T. // *Phys. Rev. C.* **100**. 054326 (2019).
- [15] Vidaña I. // *Nucl. Phys. A.* **1032**. 122625 (2023).
- [16] Wolfgruber T., Knöll M., Roth R. // *arXiv:2310.05256*

- [nucl-th]. (2023).
- [17] Knöll M., Wolfgruber T., Agel M.L. et al. // *Phys. Lett. B.* **839**. 137781 (2023).
- [18] Stoks V.G.J., Klomp R.A.M., Terheggen C.P.F., de Swart J.J. // *Phys. Rev. C.* **49**. 2950 (1994).
- [19] Шарыпов Р.Э., Мазур А.И., Широков А.М. Материалы региональной научно-практической конференции «ТОГУ-СТАРТ: Фундаментальные и прикладные исследования молодых». 2023. С. 233.
- [20] Le X., Wang N., Jiang X. // *Nucl. Phys. A.* **1038**. 122707 (2023).
- [21] Ruder S. // arXiv:1609.04747 [cs.LG]. (2023).
- [22] Haji S., Abdulazeez A. // *PalArch's J. Archaeology of Egypt / Egyptology.* **18**, N 4. 2715 (2021).
- [23] Akiba T., Sano S., Yanase T. et al. *Optuna: A Next-generation Hyperparameter Optimization Framework.* Proc. 25th International Conference on Knowledge Discovery and Data Mining. 2019.
- [24] Kingma D.P., Ba J. // arXiv:1412.6980v9 [cs.LG]. (2017).
- [25] Smith L. // arXiv:1506.01186 [cs.CV]. (2017).
- [26] TensorFlow Addons Optimizers: Cyclical Learning Rate. www.tensorflow.org/addons/tutorials/optimizers_cyclicallearningrate. 2024.
- [27] Chollet F. et al. // Keras. <https://github.com/fchollet/keras>. 2015.
- [28] Abadi M. et al. // TensorFlow: Large-scale machine learning on heterogeneous systems. <https://tensorflow.org>. 2015.
- [29] TensorFlow Addons. <https://github.com/tensorflow/addons>. 2024.

Machine Learning in the Problem of Extrapolating Variational Calculations in Nuclear Physics

A. I. Mazur¹, R. E. Sharypov^{1,a}, A. M. Shirokov²

¹*Pacific National University, Khabarovsk 680035, Russia*

²*Skobel'syn Institute of Nuclear Physics, Lomonosov Moscow State University
Moscow 119991, Russia*

E-mail: ^asharypov@pnu.edu.ru

A modified machine learning method is proposed, utilizing an ensemble of artificial neural networks for the extrapolation of energies obtained in variational calculations, specifically in the No-core Shell Model (NCSM), to the case of the infinite basis. A new neural network topology is employed, and criteria for selecting both the data used for training and the trained neural networks for statistical analysis of the results are formulated. The approach is tested by extrapolating the deuteron ground state energy in calculations with the Nijmegen II NN interaction and provides statistically significant results. This technique is applied to obtain extrapolated ground state energies of ${}^6\text{Li}$ and ${}^6\text{He}$ nuclei based on the NCSM calculations with Daejeon16 NN interaction.

PACS: 07.05.Mh, 21.10.Dr, 21.60.De.

Keywords: nuclear shell model, *ab initio* approaches, machine learning, extrapolation of variational calculations.

Received 16 February 2024.

English version: *Moscow University Physics Bulletin.* 2024. **79**, No. . Pp. .

Сведения об авторах

1. Мазур Александр Игоревич — канд. физ.-мат. наук, доцент, вед. науч. сотрудник Тихоокеанского государственного университета; e-mail: amazur@pnu.edu.ru.
2. Шарыпов Роман Эдуардович — аспирант Тихоокеанского государственного университета; e-mail: sharypov@pnu.edu.ru.
3. Широков Андрей Михайлович — канд. физ.-мат. наук, ст. науч. сотрудник НИИЯФ МГУ; e-mail: shirokov@nucl-th.sinp.msu.ru.