



Воронежский государственный университет,
факультет компьютерных наук

ОСОБЕННОСТИ ОБУЧЕНИЯ НЕЙРОННОЙ
СЕТИ БЕХЛЕРА-ПАРРИНЕЛЛО НА
РЕЗУЛЬТАТАХ КВАНТОВОЙ МОЛЕКУЛЯРНО-
ДИНАМИЧЕСКОЙ МОДЕЛИ КРИСТАЛЛА
ЛИТИЯ

Потапов Артем Геннадьевич,
Романов Александр Викторович

Москва 2025

ВВЕДЕНИЕ

КВАНТОВАЯ
ХИМИЯ [1]



МОЛЕКУЛЯРНАЯ
ДИНАМИКА

- Необходимо знать потенциал между атомами
- Для нахождения потенциала эффективно использовать нейросеть [2-4]
- Нельзя использовать традиционные координаты

НУЖНО УЧЕСТЬ СИММЕТРИЮ

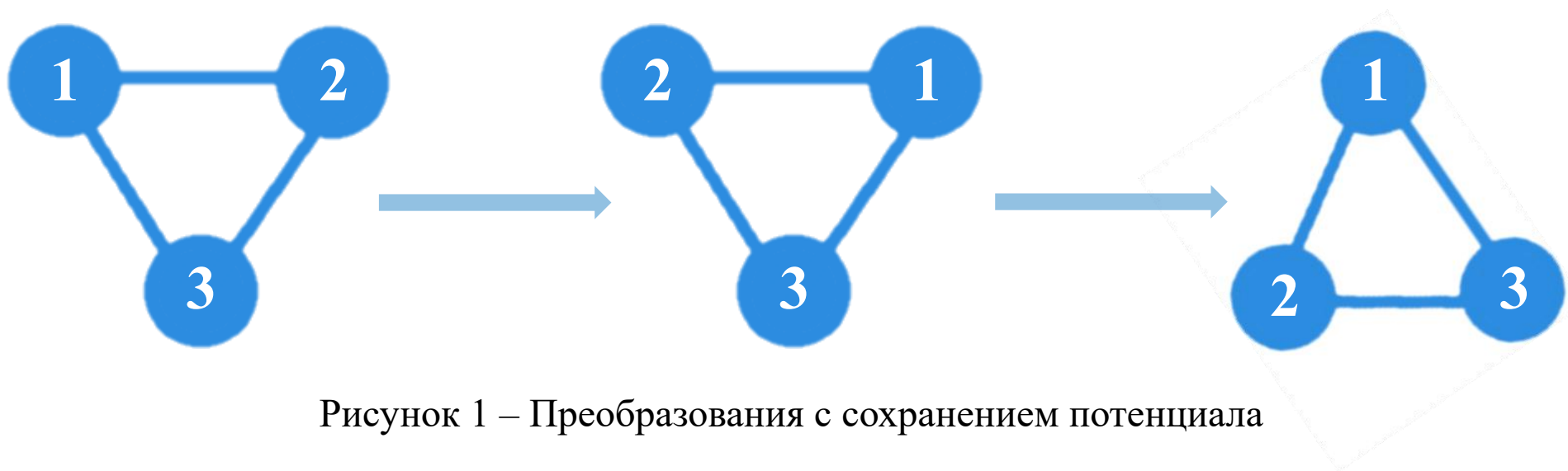


Рисунок 1 – Преобразования с сохранением потенциала

Решение: использовать симметричные функции Бехлера-Парринелло [5]

Задача: изучить способность нейросети восстанавливать межатомный потенциал с помощью значений симметричных функций



ФУНКЦИИ БЕХЛЕРА-ПАРРИНЕЛЛО

$$G_i^1 = \sum_j f_c(R_{ij}) \quad (1)$$

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}) \quad (2)$$

$$G_i^3 = \sum_j \cos(\kappa R_{ij}) \cdot f_c(R_{ij}) \quad (3)$$

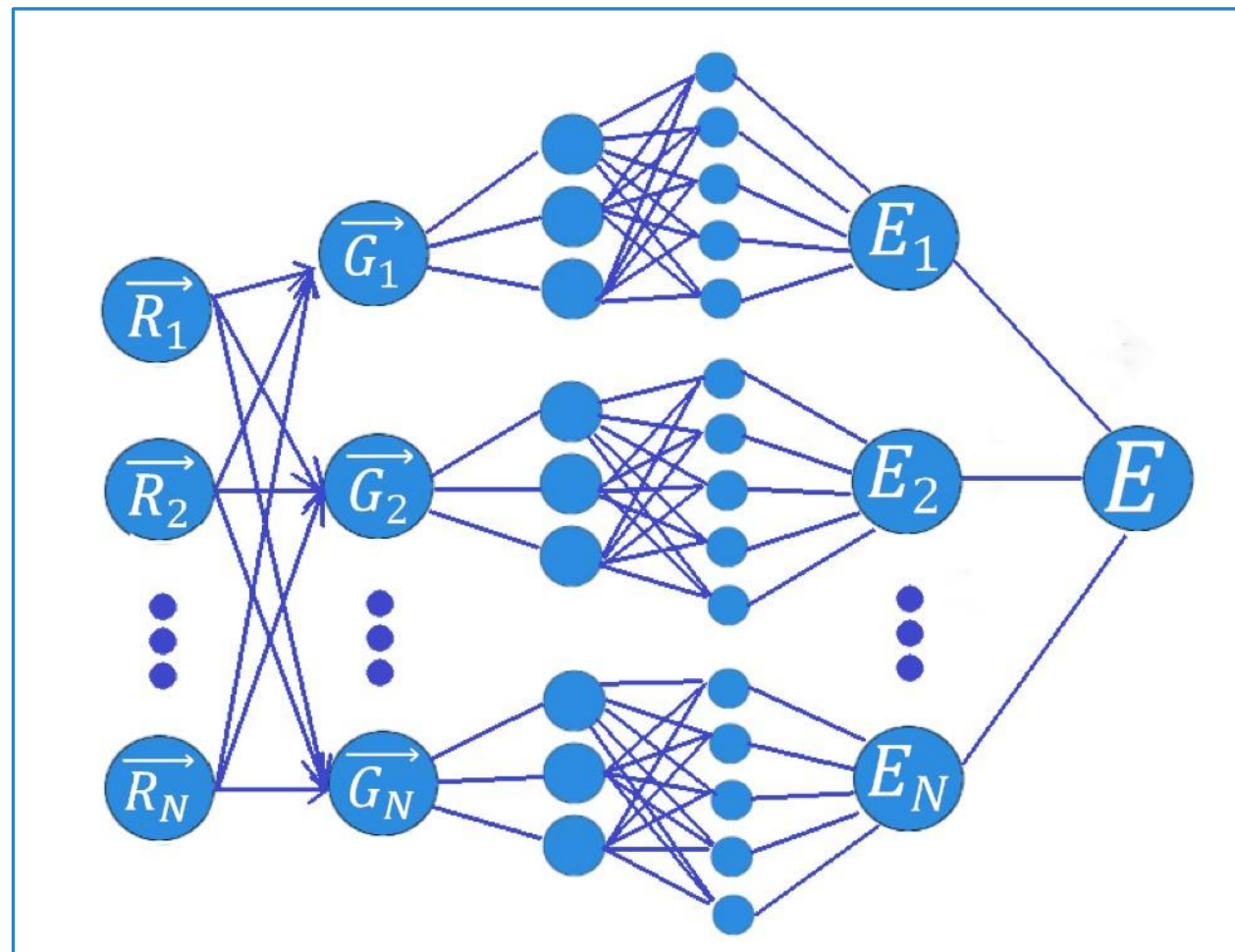
$$G_i^4 = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}) \quad (4)$$

$$G_i^5 = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \quad (5)$$

$$f_c(R_{ij}) = \begin{cases} 0,5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right], & R_{ij} \leq R_c \\ 0, & R_{ij} > R_c \end{cases} \quad (6)$$

R_{ij} — расстояние между атомами,
 R_c — радиус отсечки,
 f_c — функция отсечки,
 i — зафиксированный атом,
 η — ширина гауссиан,
 R_s — радиус сдвига центра гауссиан,
 κ — настраивает период косинуса,
 ζ — угловое разрешение,
 λ — ориентация графика $G(\theta)$,
 θ — угол между тремя атомами (в рад.)

НЕЙРОННАЯ СЕТЬ БЕХЛЕРА-ПАРРИНЕЛЛО



R – расстояния между атомами
 G – значения симметричных функций
 E – энергия системы, N – число атомов

НЕОБХОДИМОСТЬ НОРМАЛИЗАЦИИ

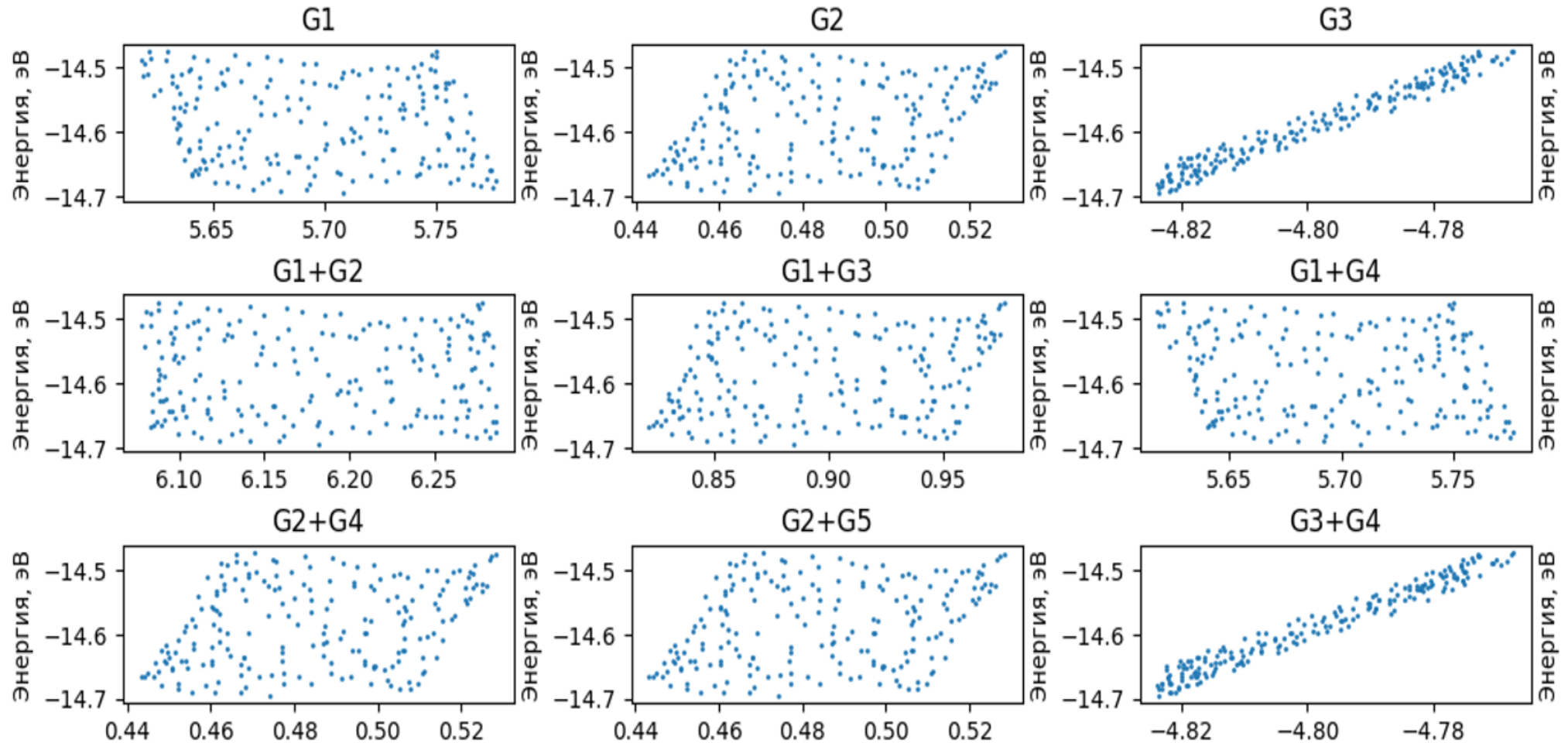


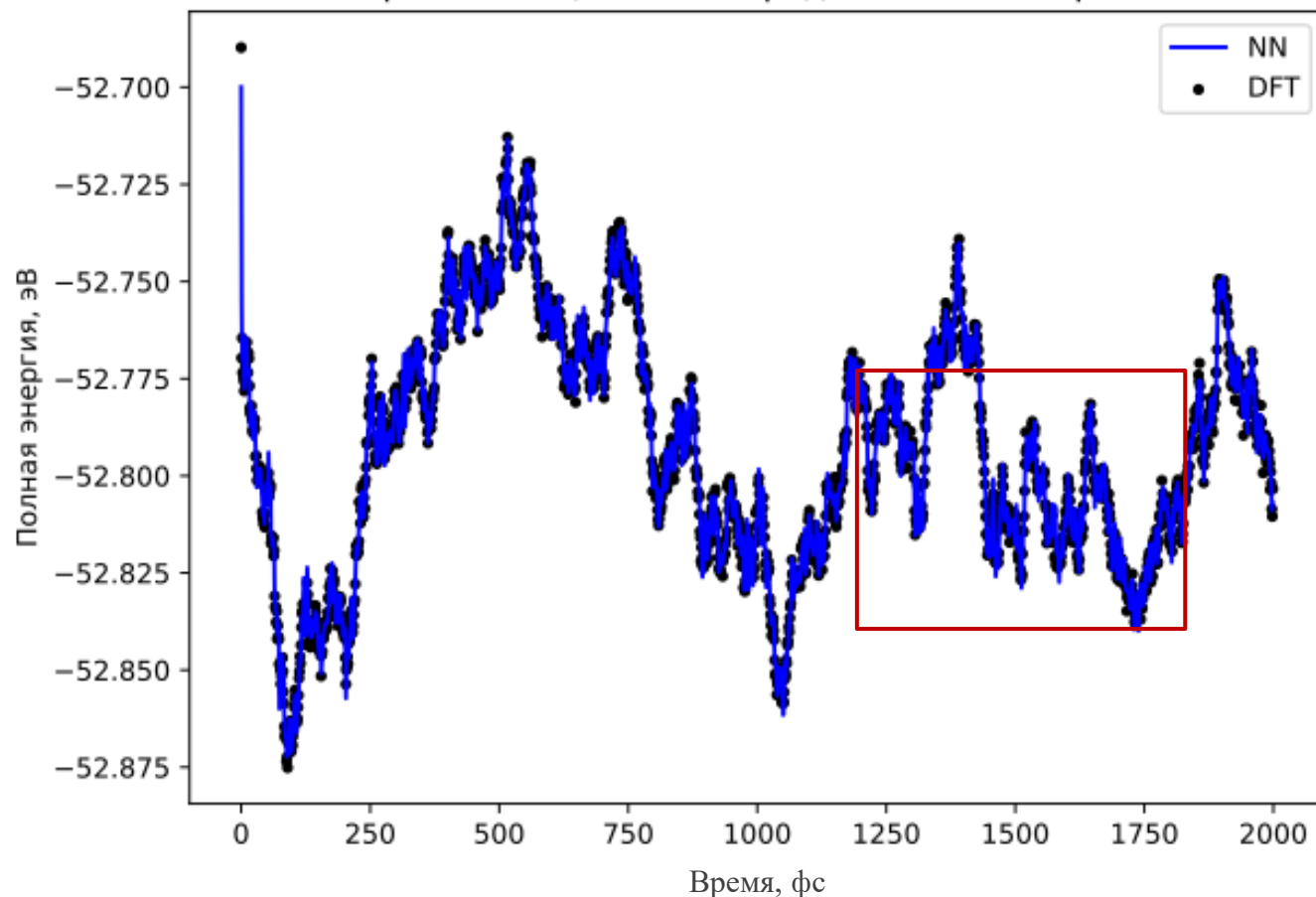
Рисунок 4 – Графики $E(G)$ для некоторых возможных комбинаций G

$$|G^1| > |G^3| > |G^2| \gg |G^5| > |G^4|$$

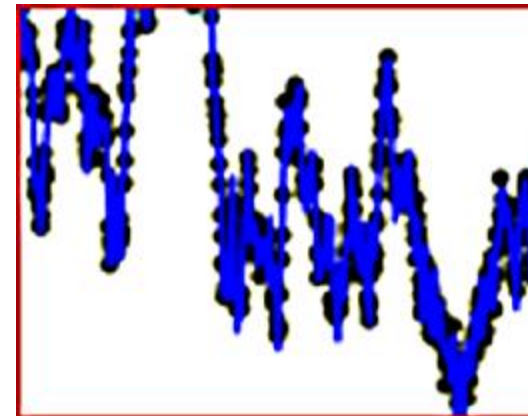
ВЛИЯНИЕ РАДИУСА ОТСЕЧКИ НА ТОЧНОСТЬ ПРЕДСКАЗАНИЯ НЕЙРОСЕТИ В ПРОГРАММЕ FORTNET [7]

Li_{27} , 2000 фемтосекунд

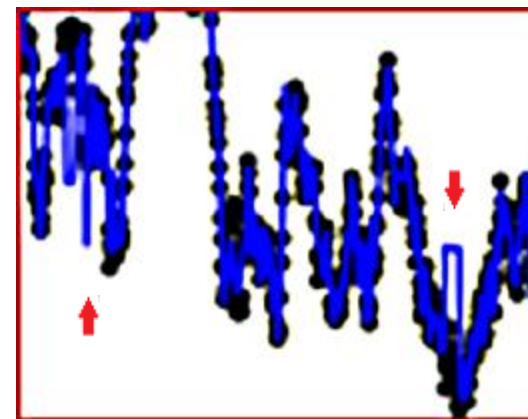
Сравнение целевой и предсказанной энергий



$R_c = 20 \text{ \AA}$



$R_c = 8 \text{ \AA}$



СРАВНИТЕЛЬНАЯ ТАБЛИЦА ВЛИЯНИЯ РАДИУСА ОТСЕЧКИ НА КАЧЕСТВО ОБУЧЕНИЯ НЕЙРОСЕТИ

Радиус отсечки, Å	Количество эпох обучения	Значение MSE	Время обучения
Li27; 2000 фс; 27/27 нейронов; 25g2+25g5			
8	4592	1.119942E-05	12 м. 34 с.
20	4283	5.951122E-06	12 м. 11 с.
Li27; 2000 фс; 27/27 нейронов; 12g1+12g4			
8	11496	2.313880E-05	23 м. 49 с.
20	20000	1.889200E-04	39 м. 55 с.
Li64; 250 фс; 32/32 нейрона; 25g2+25g5			
6	2918	4.877322E-06	5 м. 49 с.
12	1759	5.302363E-06	4 м. 31 с.
16	1848	3.759548E-06	4 м. 55 с.
16.5	2022	3.463652E-06	5 м. 13 с.
16.75	1782	3.270775E-06	4 м. 46 с.
>=17	1	NULL	0 м. 51 с.

Таблица 1 – Влияние радиуса отсечки на скорость и точность аппроксимации энергии программой Fortnet

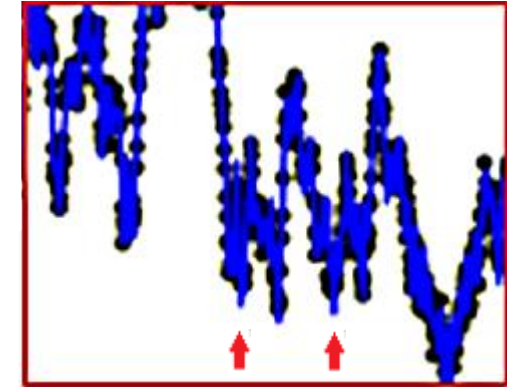
- Для $G^2 + G^5$ наилучшая точность достигается при максимальном радиусе отсечки
- Для $G^1 + G^4$ и других комбинаций подобная закономерность отсутствует
- Наибольшая скорость вычислений достигается при радиусе отсечки, достаточном для охвата каждого атома (12 Å для Li₆₄)
- Существует $R_{\max}$, для которого расчёты не производятся вследствие слишком малых значений G после нормализации
- Все расчёты проводились на 12 ядрах Intel(R) Xeon(R) CPU E5-2680 v3 2.50 GHz

ВЛИЯНИЕ ЧИСЛА СИММЕТРИЧНЫХ ФУНКЦИЙ НА ТОЧНОСТЬ ПРЕДСКАЗАНИЯ НЕЙРОСЕТИ ДЛЯ 27 АТОМОВ

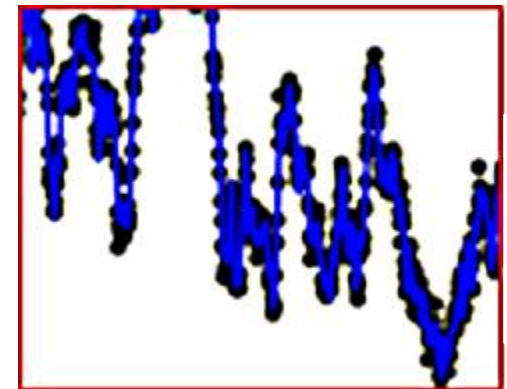
Li_{27} , 2000 фемтосекунд, 27/27 нейронов



$25G^2+25G^5$



$12G^2+12G^5$

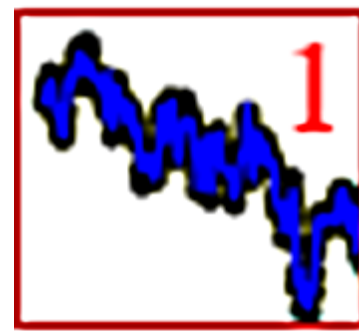


ВЛИЯНИЕ ЧИСЛА СИММЕТРИЧНЫХ ФУНКЦИЙ НА ТОЧНОСТЬ ПРЕДСКАЗАНИЯ НЕЙРОСЕТИ ДЛЯ 64 АТОМОВ

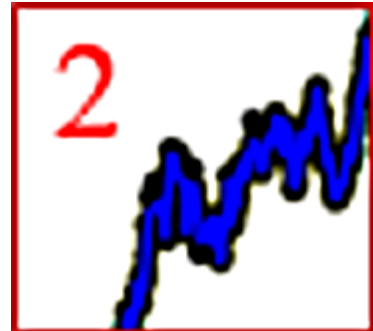
Li₆₄, 1000 фемтосекунд, 32/32 нейрона



$25G^2+25G^5$



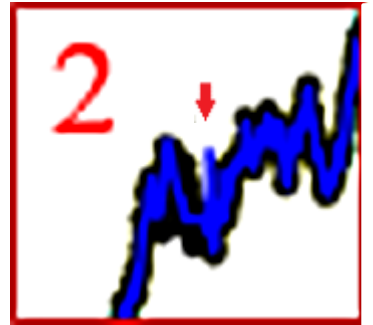
$25G^2+25G^5$



$75G^2+75G^5$



$50G^2+50G^5$



СРАВНИТЕЛЬНАЯ ТАБЛИЦА ВЛИЯНИЯ ЧИСЛА СИММЕТРИЧНЫХ ФУНКЦИЙ НА КАЧЕСТВО ОБУЧЕНИЯ НЕЙРОСЕТИ

Комбинация функций	Количество эпох обучения	Значение MSE	Время обучения
Li27; 2000 фс; 27/27 нейронов; радиус отсечки 12 Å			
6g2+6g5	16507	9.264668E-06	32 м. 52 с.
12g2+12g5	5173	5.467964E-06	11 м. 55 с.
25g2+25g5	4283	5.951122E-06	12 м. 12 с.
Li64; 1000 фс; 32/32 нейрона; радиус отсечки 12 Å			
25g2+25g5	3029	4.479973E-06	25 м. 28 с.
50g2+50g5	2998	4.386149E-06	41 м. 45 с.
75g2+75g5	3567	4.747073E-06	62 м. 40 с.
Li64; 2500 фс; 32/32 нейрона; радиус отсечки 12 Å			
25g2+25g5	3392	1.298785E-05	70 м. 21 с.
50g2+50g5	9341	1.543720E-05	301 м. 13 с.
75g2+75g5	8183	3.567265E-05	362 м. 40 с.

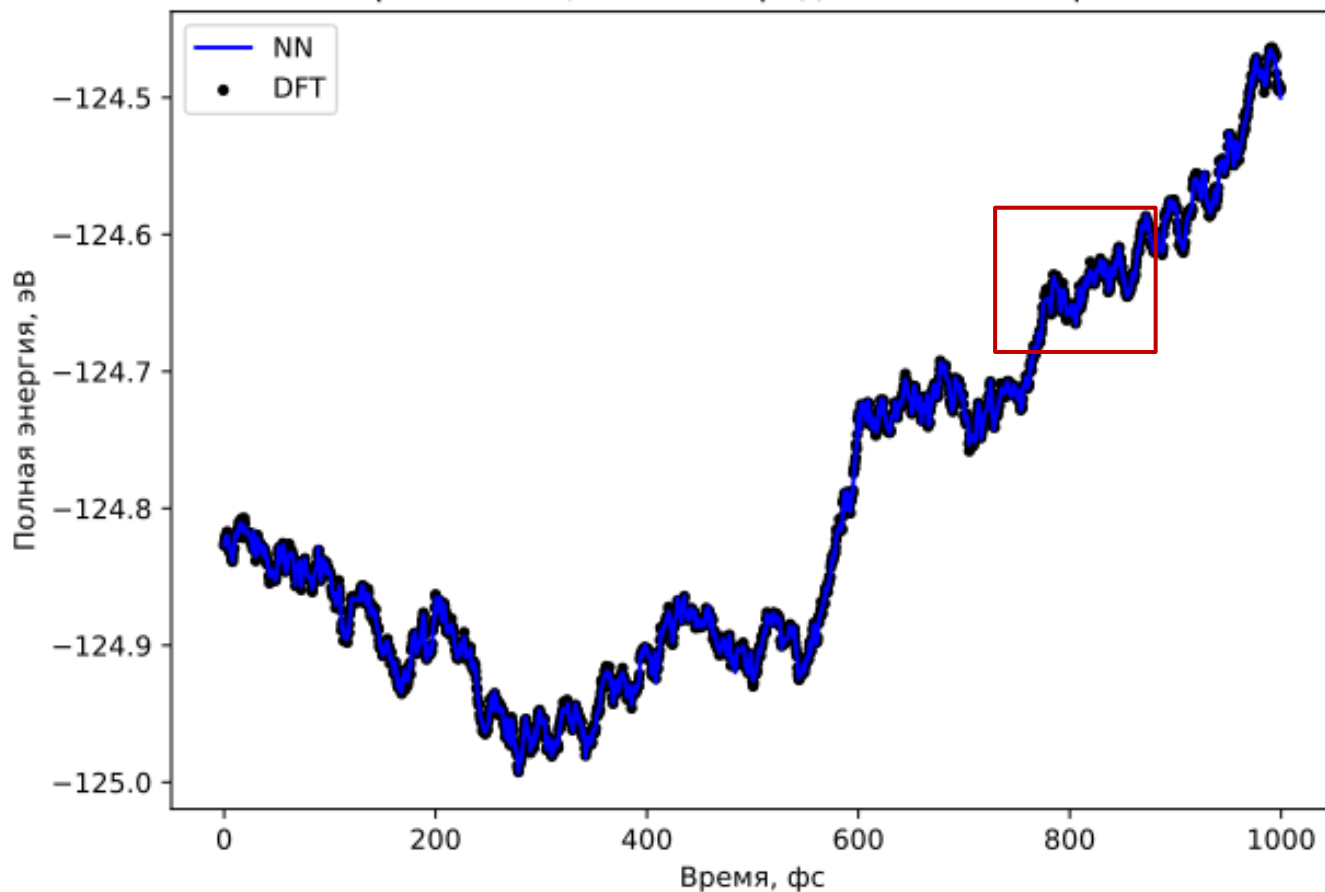
- Для **27 атомов** достаточно 24 функций
- С увеличением **числа атомов** требуется **большее** число функций при одинаковом размере датасета
- Увеличение датасета **не приводит** к необходимости увеличивать число функций
- Визуальный анализ локальных всплесков **может не совпадать** с фактически получаемой глобальной точностью

Таблица 2 – Влияние числа симметричных функций на скорость и точность аппроксимации энергии программой Fortnet

ВЛИЯНИЕ ЧИСЛА НЕЙРОНОВ СКРЫТЫХ СЛОЁВ НА ТОЧНОСТЬ ПРЕДСКАЗАНИЯ НЕЙРОСЕТИ В ПРОГРАММЕ FORTNET

Li_{64} , 1000 фемтосекунд $25G^2+25G^5$

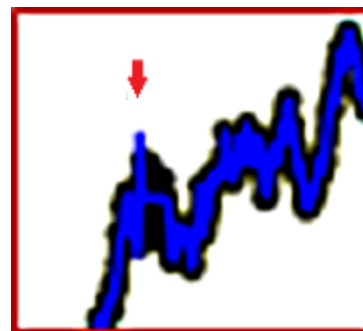
Сравнение целевой и предсказанной энергий



32/32 нейрона



64/64 нейрона



СРАВНИТЕЛЬНАЯ ТАБЛИЦА ВЛИЯНИЯ ЧИСЛА НЕЙРОНОВ СКРЫТЫХ СЛОЁВ НА КАЧЕСТВО ОБУЧЕНИЯ НЕЙРОСЕТИ

Число нейронов скрытых слоёв	Количество эпох обучения	Значение MSE	Время обучения
Li27; 2000 фс; 25g2+25g5; радиус отсечки 20 Å			
6/6	13676	4.774006E-05	13 м. 45 с.
13/13	5683	1.436364E-05	11 м. 01 с.
27/13	4221	5.811360E-06	10 м. 28 с.
27/27	4283	5.951122E-06	12 м. 12 с.
Li64; 1000 фс; 25g2+25g5; радиус отсечки 12 Å			
32/32	3029	4.479973E-06	25 м. 28 с.
64/64	3916	4.796957E-06	86 м. 31 с.
Li64; 2500 фс; 25g2+25g5; радиус отсечки 12 Å			
32/32	3392	1.298785E-05	70 м. 21 с.
64/64	4508	1.181917E-05	235 м. 16 с.

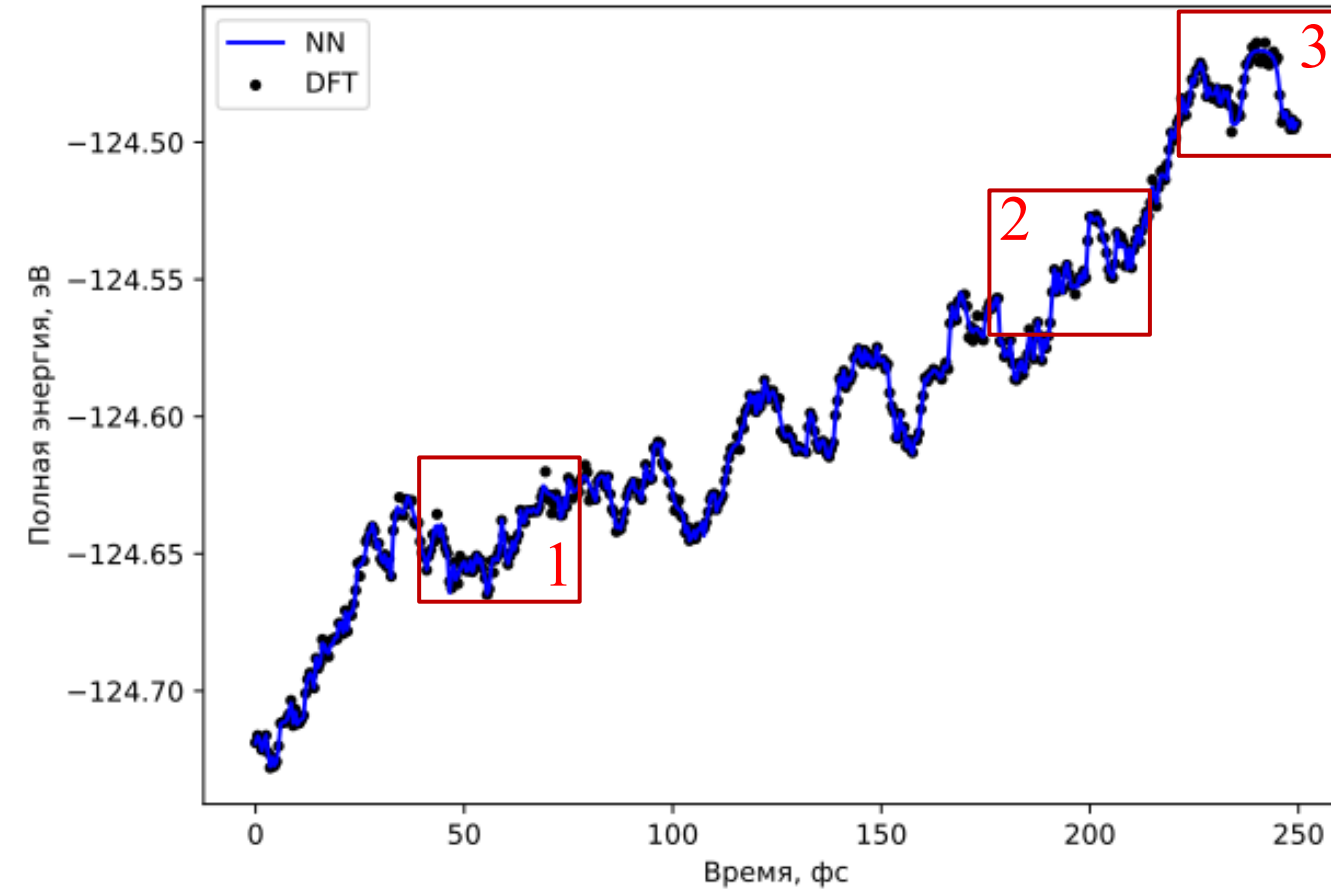
- Для 27 атомов достаточно комбинации **27/13 нейронов**
- С увеличением числа атомов **не возникает** необходимость существенно увеличивать число нейронов
- Увеличение размера датасета **приводит** к необходимости **увеличить** число нейронов
- Использование большего числа нейронов требует большего **времени обучения**

Таблица 3 – Влияние числа симметричных функций на скорость и точность аппроксимации энергии программой Fortnet

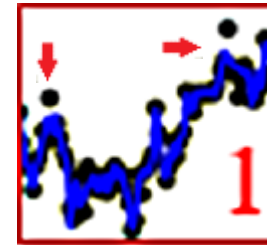
ВЛИЯНИЕ ВИДА СИММЕТРИЧНЫХ ФУНКЦИЙ НА ТОЧНОСТЬ ПРЕДСКАЗАНИЯ НЕЙРОСЕТИ В ПРОГРАММЕ FORTNET

Li_{64} , 250 фемтосекунд, 32/32 нейрона

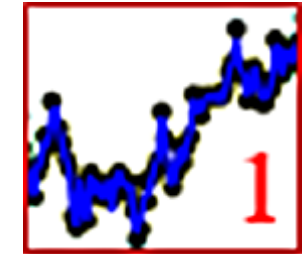
Сравнение целевой и предсказанной энергий



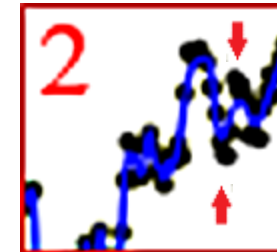
$25G^2+25G^4$



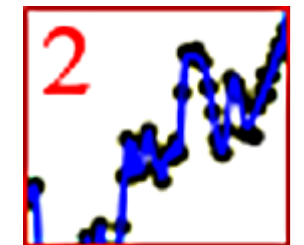
$25G^2+12G^4+12G^5$



$25G^2+12G^4+12G^5$



$25G^2+12G^4+12G^5 (*)$



$25G^2+12G^4+12G^5$



$25G^2+16G^4+8G^5$



*порог сходимости уменьшен до $1e-09$

СРАВНИТЕЛЬНАЯ ТАБЛИЦА ВЛИЯНИЯ ВИДА СИММЕТРИЧНЫХ ФУНКЦИЙ НА КАЧЕСТВО ОБУЧЕНИЯ НЕЙРОСЕТИ

Комбинация функций	Количество эпох обучения	Значение MSE	Время обучения
Li64; 250 фс; 32/32 нейрона; 50 функций; R _c = 12 Å			
12g1+12g2+12g4+12g5	3941	3.794943E-06	8 м. 40 с.
12g1+2g2+25g4	2076	3.309701E-06	5 м. 00 с.
12g1+12g2+25g5	4836	3.807081E-06	10 м. 28 с.
12g1+12g3+12g4+12g5	7045	5.471005E-06	14 м. 32 с.
12g1+12g3+25g4	6308	4.171834E-06	13 м. 40 с.
12g1+12g3+25g5	5533	6.690832E-06	11 м. 28 с.
12g2+12g3+12g4+12g5	3526	3.727510E-06	7 м. 48 с.
12g2+12g3+25g4	2999	3.992554E-06	6 м. 55 с.
12g2+12g3+25g5	3732	5.266695E-06	8 м. 12 с.
25g1+12g4+12g5	20000	1.310946E-05	40 м. 21 с.
25g1+25g4	19999	1.067284E-04	39 м. 51 с.
25g1+25g5	19999	1.858414E-04	39 м. 49 с.
25g2+12g4+12g5	1591	2.410607E-06	3 м. 56 с.
25g2+12g4+12g5	6452	6.871266E-07	13 м. 25 с.
25g2+12g4+12g5	50000	2.896467E-07	244 м. 48 с.
25g2+16g4+8g5	1660	1.687497E-06	4 м. 04 с.
25g2+18g4+12g5	1831	1.051187E-05	4 м. 40 с.
25g2+25g4	2207	2.294304E-06	5 м. 27 с.
25g2+25g5	2000	4.558489E-06	4 м. 43 с.
25g3+12g4+12g5	7231	7.038860E-06	15 м. 15 с.
25g3+25g4	19999	1.035068E-04	40 м. 03 с.
25g3+25g5	17115	5.151634E-06	35 м. 25 с.
8g1+8g2+8g3+12g4+12g5	3339	9.804373E-06	7 м. 25 с.
8g1+8g2+8g3+25g4	3339	1.130260E-05	7 м. 43 с.
8g1+8g2+8g3+25g5	4831	7.278316E-06	10 м. 31 с.

- Среди радиальных функций самую высокую точность демонстрирует G^2 . Самую низкую - G^1
- Среди угловых функций комбинация $G^4 + G^5$ приводит, в среднем, к меньшему значению ошибки, чем только G^4 или только G^5 . При этом комбинировать их следует не поровну, а в пропорции **2:1**
- При прочих равных лучший результат показала комбинация **$25G^2 + 16G^4 + 8G^5$**
- Наименее точные результаты зафиксированы у комбинации **$25G^1 + 25G^5$**
- Введение более строгих ограничений на сходимость может привести к **более точным результатам**, но требует **больше времени** на обучение

ЗАКЛЮЧЕНИЕ

- С помощью нейронной сети возможно аппроксимировать значение энергии системы в каждый момент времени, основываясь **только на данных о взаимных межатомных расстояниях и валентных углах**
- Совокупность значений функций Бехлера-Парринелло представляет собой **многомерное признаковое пространство**, которое невозможно визуализировать, но которое успешно могут обработать нейросети (в частности, сеть Бехлера-Парринелло, реализованная в программе Fortnet)
- Наибольшая точность вычислений достигается при **максимально возможном** радиусе отсечки. При этом слишком большие значения **недопустимы**, поскольку приводят к NULL-значениям симметричных функций
- С увеличением числа атомов требуется **большее число функций** при одинаковом размере датасета
- Увеличение размера датасета приводит к необходимости **увеличить число нейронов** в скрытых слоях
- При прочих равных лучший результат показала комбинация $G^2 + G^4 + G^5$, взятая в пропорции **25:16:8**
- В дальнейшем планируется расширять диапазон исходных данных и использовать обученную нейросеть в моделировании методом классической молекулярной динамики

ЛИТЕРАТУРА

- [1] Marx D., Hutter J. Ab initio Molecular Dynamics: Basic Theory and Advanced Methods // Cambridge University Press. 2009. Vol. 63, No. 3. 567 p. DOI: 10.1017/CBO9780511609633.
- [2] Behler J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review // International Journal of Quantum Chemistry. 2015. Vol. 115, No. 16. 19 p. DOI: 10.1002/qua.24890.
- [3] Kyuhyun L., Dongsun Y., Wonseok J., et al. SIMPLE-NN: An efficient package for training and executing neural-network interatomic potentials // Computer Physics Communications. 2019. Vol. 242. P. 95–103. DOI: 10.1016/j.cpc.2019.04.014.
- [4] Blank T.B., Brown S.D., Calhoun A.W., et al. Neural network models of potential energy surfaces // The Journal of Chemical Physics. 1995. Vol. 103, No. 10. P. 4129–4137. DOI: 10.1063/1.469597.
- [5] Behler J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials // The Journal of Chemical Physics. 2011. Vol. 134. No. 3. 14 p. DOI: 10.1063/1.3553717.
- [6] Потапов А.Г., Романов А.В. Моделирование межатомного потенциала по результатам квантово-химического расчёта с использованием симметричных функций Бехлера-Парринелло // Труды молодых ученых факультета компьютерных наук ВГУ. Том Выпуск 4. Воронеж, 2024. С. 215–222.
- [7] Van der Heide T., Kullgren J., Broqvist P., et al. Fortnet, a software package for training Behler Parrinello neural networks // Computer Physics Communications. 2023. Vol. 284. 13 p. DOI: 10.5281/zenodo.5969907.