

01

## Машинообученный потенциал Li–C для наноматериалов

© С.А. Созыкин

Южно-Уральский государственный университет  
(Национальный исследовательский университет),  
454080 Челябинск, Россия  
e-mail: sozykinsa@susu.ru

Поступило в Редакцию 2 октября 2024 г.

В окончательной редакции 2 октября 2024 г.

Принято к публикации 2 октября 2024 г.

Рассмотрена разработка потенциалов межчастичного взаимодействия двух типов для атомистического моделирования комплексов углеродных наноматериалов с литием. Первый потенциал построен методом гауссовой аппроксимации, второй — в рамках подхода глубокого обучения. Эти потенциалы обучались на результатах моделирования методом функционала электронной плотности и обеспечивали точность, близкую к точности этого метода при существенно меньших требованиях к вычислительным ресурсам. Наборы данных содержали более 8000 структур приблизительно из 100 атомов. Конкретная структура с вероятностью 90 % помещалась в обучающую выборку, в противном случае — в валидационную. Полученные потенциалы позволили точно воспроизвести энергии комплексов и силы, действующие на атомы. Время расчета линейно возрастает с количеством атомов в модели и может изменяться на несколько порядков в зависимости от типа потенциала и используемого аппаратного обеспечения. Потенциал, полученный методом глубокого обучения, кажется многообещающим для реалистичного и точного моделирования лития на поверхности углеродных нанотрубок и различных графеноподобных структур при температурах до 450 К.

**Ключевые слова:** молекулярная динамика, глубокое обучение, регрессия, литий, углеродные наноматериалы.

DOI: 10.61011/JTF.2025.03.59849.292-24

### Введение

Взрывной рост потребности в портативных источниках энергии вызвал активное исследование углеродных наноструктур в качестве материала анода литиевых батарей [1]. Эксперимент показывает, что использование в аноде батареи многослойного графена [2] или углеродных нанотрубок (УНТ) [3] вместо графита позволяет увеличить ее обратимую литиевую емкость. В случае вертикально ориентированных УНТ [4] путь диффузии лития вглубь анода не вызывает сомнений в отличие от графена, который может иметь чешуйчатую структуру, способствующую адсорбции лития по границам чешуек. Такое неравномерное распределение лития по поверхности может приводить к малой чувствительности процесса интеркаляции к концентрации дефектов графена [5].

Для понимания процессов, протекающих в аноде литиевой батареи в ходе процессов разрядки/зарядки, необходимы реалистичные модели — цифровые двойники таких анодов. Основывать такие модели исключительно на квантово-механических методах не представляется возможным из-за сложности структуры анода и, следовательно, большого размера расчетной ячейки. Например, предложенный в работе [6] перспективный материал анода батареи представляет собой графен, выращенный на ковалентно связанных с графеном УНТ. В работе [7] описана перспективность применения выращенных на графене УНТ в качестве анодов литий-серных аккумуля-

торов. В работе [8] сообщается о достижении выдающихся характеристик батареи при использовании в качестве анода композиционного материала, содержащего УНТ. Эти примеры показывают, что структура перспективных материалов много сложнее, чем массивы нанотрубок или стопки листов графена. Очевидно, что такие материалы содержат углерод с различной гибридизацией и построены они не только из шестиугольников углерода. Таким образом, актуальной задачей становится изучение поведения лития вблизи поверхности сложным образом сочлененных углеродных наноматериалов и развитие соответствующих методов моделирования.

Распространенным теоретическим подходом к изучению адсорбции лития на углеродных поверхностях является теория функционала электронной плотности. У нее есть неоспоримое преимущество: практически отсутствует необходимость в априорных сведениях об исследуемой системе, а достоверность получаемых результатов ограничена только доступными вычислительными ресурсами. Последнее обстоятельство заметно ограничивает размер модели, доступный для изучения. Альтернативой выступает классическая молекулярная динамика с межчастичными потенциалами взаимодействия, зависящими от нескольких параметров. Параметры эти можно подбирать из разных соображений, но получаемый потенциал оказывается применим только для узкого класса объектов (близких к тем системам, под которые проводилась подгонка параметров).

**Таблица 1.** Параметры  $c$  ячейки АВ-графита и длины связей С–С,  $L_{C-C}$ , полученные с использованием разных обменно-корреляционных функционалов (EX)

EX	$L_{C-C}$ , Å	$c$ , Å	EX	$L_{C-C}$ , Å	$c$ , Å
LDA	1.412	6.588	vdW-DF-cx of Berland and Hyldgaard	1.426	6.833
PBE	1.424	7.852	SCAN+rVV10 of Peng	1.416	6.627
PBE-D3	1.423	6.905	rVV10 of Sabatini	1.424	6.348
optB88-vdW	1.424	6.638	r <sup>2</sup> SCAN+rVV10 of Ning	1.418	6.620
optB86b-vdW	1.423	6.566	Original vdW-DF of Dion	1.430	7.076
rev-vdW-DF2	1.423	6.595	optPBE-vdW of Klimeš	1.427	6.792
vdW-DF2 of Lee	1.429	6.943	Эксперимент	1.418	6.708

Качественный скачок в улучшении предсказаний „классического“ моделирования может быть получен с помощью методов искусственного интеллекта, где потенциал строится по данным расчетов методом теории функционала электронной плотности (DFT), но не в виде явной функции, а в виде нейронной сети. Это очень бурно развивающееся направление исследований, открывающее возможность проведения моделирования методом классической молекулярной динамики с точностью, не уступающей методу неэмпирической молекулярной динамики. Безусловно, DFT успешно используют для подгонки параметров потенциалов и без использования искусственных нейронных сетей. Примером выступают потенциалы GAP (Gaussian Approximation Potentials) [9].

На сегодняшний день известны несколько реализаций машиннообучаемых потенциалов, ставящих своей целью заменить для масштабных и длительных расчетов метод DFT. Наиболее распространенные из них — MLIP (Machine Learning Interatomic Potentials) [10], DP (Deep Potentials) [11] и CHGNet (Crystal Hamiltonian Graph neural Network) [12]. Все перечисленных модели были разработаны за последнее десятилетие и активно развиваются в настоящее время. Все они могут быть использованы совместно с программой LAMMPS для моделирования методом классической молекулярной динамики.

Несомненный успех в разработке связанных с атомистическим моделированием методов машинного обучения был бы еще более заметен научной общественности, если бы было доступно большее количество готовых к использованию межчастичных потенциалов. Например, есть универсальные потенциалы для углерода [13,14], для двумерных углеродных структур [15], для углеродных структур с учетом взаимодействия Ван-дер-Ваальса [16,17], сообщается о разработке DP-потенциала [18] (сам потенциал недоступен). Для систем лития и углерода есть только потенциал [19] (используется совместно с [13]).

Данное исследование направлено на построение DP-потенциала для описания взаимодействия лития с поверхностью УНТ и графеноподобных структур. Потенциалы и примеры входных файлов для программы

LAMMPS с их использованием приведены в дополнительных материалах к настоящей работе.

## 1. Метод

Большинство моделей машинного обучения, включая DP и GAP, предполагают, что потенциальная энергия системы может быть представлена суммой вкладов энергий каждого атома, описывающих многочастичные взаимодействия между атомами в пределах некоторого фиксированного расстояния. Этот подход удобен с вычислительной точки зрения, поскольку его сложность линейно возрастает с размером системы. Во многих случаях пренебрежение дальнедействующим кулоновским взаимодействием не приводит к большим погрешностям. Отчасти это связано с тем, что обучение моделей проводится на результатах моделирования DFT, учитывающих электростатическое взаимодействие. Для тех редких случаев, когда такого косвенного учета оказывается недостаточно, существуют обобщения методов машинного обучения для точного учета дальнедействующих электростатических взаимодействий [20].

Следствием указанной особенности машиннообучаемых потенциалов является наличие некоторого радиуса отсечки, влиянием атомов за пределами которого пренебрегают. Для исключения нефизических взаимодействий параметры трансляции должны быть больше удвоенного радиуса отсечки. При выборе параметра отсечки полезно учесть характерные масштабы взаимодействий в системе. Так, ковалентные взаимодействия С–С наблюдаются на расстояниях менее 0.16 nm. На расстояниях порядка 0.36 nm в графите наблюдаются только слабые ван-дер-ваальсовы взаимодействия. Литий в основном состоянии располагается на расстояниях около 0.25 nm от углеродной поверхности. В настоящей работе радиус отсечки выбирался равным 0.6 nm.

Все приведенные в настоящей работе результаты квантохимических расчетов были получены в программе VASP [21]. Известно, что с описанием взаимодействия слоев графита в DFT есть проблемы [22]. Параметр  $c$  ячейки АВ-графита и длина связи С–С,  $L_{C-C}$ , в нем, полученные при использовании ряда обменно-корреляционных функционалов, приведены в табл. 1.

Из таблицы видно, что большинство функционалов завьшают длину связи С–С по сравнению с экспериментальным значением. Тем не менее отклонения не превышают 0.012 Å. С параметром  $c$  решетки дело обстоит хуже: наиболее часто упоминаемый в научных статьях функционал PBE приводит к существенной переоценке данной величины. Использование эмпирической поправки PBE-D3 ситуацию исправляет, но ее эмпирический характер не позволяет быть уверенным в корректности предсказания поправки к энергии для более сложных, чем графит углеродных наноматериалов с литием. Ситуация осложняется тем, что структуры, на которых проводится обучение потенциала, должны быть далекими от равновесия: для равновесных структур практически равны нулю действующие на атомы силы, что снижает объем полезных данных для обучения потенциала. Из табл. 1 видно, что существует целый ряд функционалов, которые хорошо описывают и межатомное, и межплоскостное расстояние в графите. К сожалению, использование всех их за исключением излишне упрощенного функционала LDA сопряжено с колоссальными в сравнении с PBE вычислительными затратами, что не позволило рассматривать обсуждаемые в настоящей работе модели на имеющихся у автора вычислительных ресурсах. В связи с изложенным выше в настоящей работе расчеты проводились с использованием обменно-корреляционного функционала PBE. Расчеты проводились без учета спина с использованием PAW псевдопотенциалов. Энергия отсечки выбиралась равной 600 eV. Разбиение обратного пространства на  $k$ -точки выполнялось по схеме Монхорста–Пака с разными параметрами для одномерных ( $1 \times 1 \times 29$ ) и двумерных ( $6 \times 6 \times 1$ ) моделей, обеспечивающими сходимость по энергии не хуже 0.1 meV/atom.

Потенциалы создавались двумя независимыми инструментами: DeePMD-kit (v2.2.6) [11] и GAP [23]. Первый из них предполагает глубокое обучение (искусственные нейронные сети со скрытыми слоями), а во втором при создании потенциала используется непараметрическая регрессия, которая не предполагает закрытую функциональную форму и не пытается объяснить процесс, лежащий в основе данных, используя теоретические соображения: гибкая функция подгоняется по большим объемам данных [24]. На вход обоих программ подавались идентичные обучающие и валидационные наборы данных, содержащие координаты и тип атомов, действующие на них силы, энергии систем и вириалы (взяты с обратным знаком сумма по всем атомам системы скалярных произведений радиус-вектора атома на силу, действующую на атом).

Стоит отметить, что подготовка машинообучаемых потенциалов требует существенных вычислительных ресурсов. В случае потенциала GAP требуются вычислительные узлы с большим объемом оперативной памяти (сотни гигабайт / узел). Для обучения качественных DP-потенциалов нужны мощные видеокарты суперком-

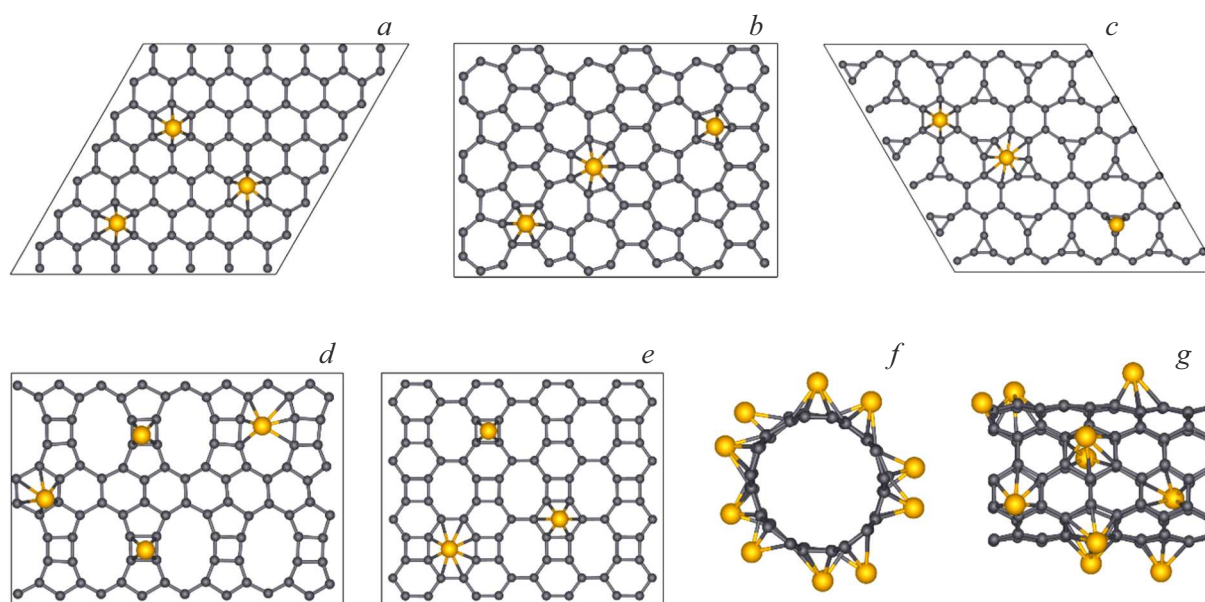
пьютерного уровня: Tesla P100, Tesla V100, Tesla A100 и т.д.

При обучении потенциала при помощи программы DeePMD для сравнения использовались дескрипторы  $se\_e2\_a$  и  $se\_e3$ . Оба этих дескриптора относятся к „Smooth Edition“ версии дескрипторов. Для дескрипторов этого типа выделяют две концентрические сферы, в центре которых расположен конкретный атом. Сфера отсечки с большим радиусом ограничивает область пространства, взаимодействие с атомами в которой учитывается при обучении потенциала. При этом взаимодействие с атомами внутри малой сферы учитывается в полной мере, а с атомами в области между сферами с весом, спадающим до нуля на сфере большого радиуса. Сокращения  $e2$  и  $e3$  указывают на то, что первый из них учитывает двухчастичные взаимодействия, а второй — взаимодействия троек атомов. Это, однако, не означает, что дескриптор  $se\_e2\_a$  учитывает лишь расстояния между атомами. Он учитывает их взаимное расположение (радиус-вектора).

Большая часть параметров дескрипторов  $se\_e2\_a$  и  $se\_e3$  совпадают. В настоящей работе значения для них так же выбирались одинаковыми. Радиус отсечки устанавливался равным 0.6 nm, при этом дескриптор плавно уменьшался, начиная с расстояния 0.4 nm. Максимальное количество рассматриваемых атомов углерода и лития в окружении данного атома устанавливалось равным 160 и 80 соответственно. Этот параметр выбирался из соображений, что количество атомов углерода, ограниченных сферой радиуса 0.6 nm, в наиболее плотной упаковке составляет около 160. Атомов лития в системе в разы меньше. Скрытая сеть содержала три слоя с размерами 50, 100 и 200. Подгоночная сеть содержала три слоя, каждый размером 250. Финальная скорость обучения составляла  $10^{-8}$ , обучение проводилось на протяжении  $10^6$  эпох. Конечные веса энергий, сил и вириала в функции потерь составляли 1.5, 1.0 и 0.2 соответственно. Кроме того, был рассмотрен вариант совместного использования дескрипторов  $se\_e2\_a$  и  $se\_e3$  с уменьшенным радиусом отсечки для  $se\_e3$ .

При создании GAP-потенциала использовались дескрипторы  $distance\_Nb$  второго и третьего порядков, дескриптор SOAP. Все параметры для этих дескрипторов были взяты из работы [19] с некоторыми исключениями для дескриптора SOAP: параметры  $n\_max$  и  $l\_max$  устанавливались равными 9, параметр отсечки был равен 4.0 nm (максимальное значение, при котором расчет запускался на четырех вычислительных узлах с 250 GB оперативной памяти на каждом). Таким образом, характеристиками структуры в пределах некоторой сферы, окружающей конкретный атом, выступали:

- 1) расстояния между атомами (двухчастичный дескриптор);
- 2) углы, образованные тройками атомов (трехчастичный дескриптор);
- 3) комплексная информация о взаимном положении атомов (многочастичный дескриптор SOAP).



**Рис. 1.** Представители классов моделей: *a* — графен+3Li, *b* — псиграфен+3Li, *c* — ирида-графен+3Li, *d* — тетра-пента-дека-гексагональный графен+4Li, *e* — бифенилен+3Li, *f, g* — УНТ(5,5)+10Li.

Дескриптор SOAP для каждого атома представляет собой вектор из сотен компонентов (тысяч для систем с несколькими элементами). Их количество зависит от нескольких параметров, грамотно подобрав которые можно без ущерба для качества обучаемого потенциала достичь существенного сокращения данных, подаваемых на вход процедуре обучения. Использование дескриптора SOAP позволяет существенно увеличить качество получаемого потенциала и в разы увеличивает требуемые для его обучения вычислительные ресурсы.

## 2. Результаты и обсуждение

### 2.1. Обучающий, валидационный и тестовый наборы

Создание базы данных структур проводилось в несколько этапов. На первом этапе в базу помещалось небольшое число (около 100) шагов *ab-initio* молекулярной динамики моделей углеродных наноструктур (нанотрубка, графен, бифенилен, ирида-графен,  $\psi$ -графен, тетра-пента-дека-гексагональный графен) с литием, примеры которых приведены на рис. 1. Модели содержали около 100 атомов углерода и до 10 атомов лития. Такие концентрации соответствуют „газу“ и „островкам“ лития на углеродной поверхности [25]. Новые структуры добавлялись согласно алгоритму DP-Gen [26] из молекулярно-динамического моделирования при температурах до 450 К. Размер элементарной ячейки поддерживался равным 2.5 нм в направлениях, перпендикулярных оси нанотрубок, и в направлении, перпендикулярном плоскости графеноподобных структур. Остальные параметры трансляции брались с соответствующего

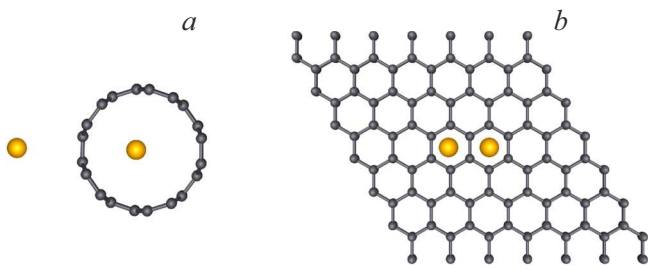
шага молекулярной динамики. Таким образом, база структур описывает не взаимодействующие со своими периодическими изображениями нанотрубки и графеноподобные объекты. После увеличения базы в несколько раз начальные структуры, полученные в методе молекулярной динамики, исключались из базы. Таким образом, все структуры в базе были отобраны DP-Gen. База данных содержала 8723 структуры, в том числе 513 структур без атома лития (63 УНТ(5,5), 50 графена, 100 бифенилена, 100 ирида-графена, 100  $\psi$ -графена и 100 тетра-пента-дека-гексагонального графена), 778 УНТ(5,5)+1Li, 328 УНТ(5,5)+5Li, 1498 УНТ(5,5)+10Li, 209 графен+1Li, 230 графен+3Li, 1262 бифенилен+3Li, 1342 ирида-графен+3Li, 1211  $\psi$ -графен+3Li, 1356 тетра-пента-дека-гексагональный графен+4Li и одна структура изолированного атома лития.

Конкретная модель с вероятностью 90% помещалась в обучающую выборку, в противном случае она попадала в валидационную выборку. Тестовую выборку формировали два набора данных:

- 1) 70 структур с атомом лития на прямой, перпендикулярной гексагону стенки нанотрубки и проходящей через его центр (рис. 2, *a*);
- 2) 30 структур с атомом лития на прямой, параллельной графену, с кратчайшими расстояниями до центров шестиугольников, соответствующих равному расстоянию от лития до графена (рис. 2, *b*).

### 2.2. Оценка качества потенциала

Среднеквадратическая ошибка (root mean squared error, RMSE) часто выступает количественной мерой ошибок, которую модель машинного обучения стремится



**Рис. 2.** Крайние положения атома лития в тестовых наборах: *a* — 1 и *b* — 2.

ся минимизировать. Для оценки точности потенциала по отношению к различным типам структур были рассчитаны RMSE на валидационной базе. Результаты для четырех потенциалов приведены в табл. 2 для каждого класса объектов. DP-e2, DP-e3 и DP-hybrid обозначают потенциалы, полученные для случаев использования дескрипторов *se\_e2\_a*, *se\_e3* и гибридного соответственно. Видно, что гибридный потенциал приводит к наименьшим ошибкам. Максимальные ошибки получаются при использовании дескриптора *se\_e3*. Оба указанных потенциала оказались многократно более требовательными к вычислительным ресурсам, чем потенциал DP-e2. Именно его будем рассматривать далее в настоящей работе. Для него в семи из десяти случаев ошибка оказывается меньше 1 meV/atom. Для трех оставшихся — меньше 2 meV/atom. Для сил RMSE — не превышает 0.115 eV/Å. Потенциал GAP обеспечивает точность, сравнимую с потенциалами DP лишь для одной системы (УНТ(5,5) с пятью атомами лития). В остальных случаях ошибки оказываются до нескольких раз выше. Это указывает на сложность получения качественных GAP-потенциалов для случая большого разнообразия моделей в обучающей и тестовых выборках. Данный факт не отражает принципиальное ограничение подхода GAP, а лишь демонстрирует ограниченность доступных автору вычислительных ресурсов.

Судя по оценкам RMSE (табл. 2), полученные потенциалы PD не уступают приведенным в литературе. Для кварцевого стекла наименьшие RMSE энергии и силы составили 4 meV/atom и 0.289 eV/Å соответственно [27]. Чуть большую RMSE (4.1 meV/atom) получали для схожих по структуре кластеров золота [28]. При попытке обучения на базе более разнообразных кластеров золота ошибка возрастала в разы. DP-потенциал, обученный на наборе из множества углеродных материалов [18], дает RMSE 17 meV/atom для валидационной выборки из структур графена. Оценки ошибок существенно меньше в случае обучения модели на однотипных структурах [29]. Таким образом, полученные в настоящей работе потенциалы DP можно рекомендовать к использованию при изучении комплексов различных углеродных поверхностей с атомами лития. Несмотря на то что оценки ошибок для разных классов моделей отличаются, все они невелики по сравнению с типичны-

ми ошибками, сообщаемыми в литературе для подобных потенциалов.

Отметим, что для случая потенциала DP судить о качестве потенциала на основе зависимости RMSE от эпохи, которую выдает программа DeepMD в процессе обучения, можно только в том случае, когда структуры, на которых проводится обучение, одинаково хорошо описываются потенциалом. Дело в том, что указанные данные рассчитываются не для всей обучающей/валидационной базы, а для случайным образом взятой выборки из них. Это приводит к тому, что указанная кривая даже после очень большого числа эпох испытывает существенные колебания.

Основным стимулом разработки потенциалов молекулярной динамики, дающих схожие с DFT результаты, выступает стремление сократить время расчета. По этой причине важно, помимо точности, обращать внимание на продолжительность расчетов. Зависимость времени выполнения одного шага молекулярной динамики в программе LAMMPS от типа потенциала (DP-e2 или GAP) и используемого аппаратного обеспечения приведена на рис. 3. Приведенные на рисунке зависимости близки к линейным. Расчеты проводились для серии моделей графена с адсорбированным литием, полученных транслированием расчетной ячейки  $C_{98}Li_3$ . Минимальная из рассмотренных систем содержала 404 атома. Тестирование проводилось на компьютерах с процессорами Intel i9 10980XE (18 ядер) и AMD EPYC 7742 (64 ядра). Расчеты с потенциалом DP дополнительно запускались на видеокартах NVIDIA Tesla P100 (16 GB видеопамати), Titan V (12 GB видеопамати) и Tesla A100 (80 GB видеопамати). Из рис. 3 видно, что расчеты с использованием потенциала GAP в десятки раз более продолжительны, чем расчеты с потенциалом DP на том же процессоре. По сравнению с запуском расчета на графическом ускорителе, проигрыш во времени исчисляется уже сотнями раз. На фоне таких больших изменений времени расчета отличия во времени расчета на разных видеокартах и на разных процессорах не кажутся существенными. Из-за большой требовательности к вычислительным ресурсам тестирование потенциала GAP проводилось только на относительно маленьких системах — до 25 856 атомов. Несмотря на относительно низкие требования к вычислительным ресурсам для потенциала DP, в случае использования графических ускорителей максимальный размер системы, доступной для моделирования, оказывается ограничен размером видеопамати, который, как правило, существенно меньше оперативной памяти центрального процессора. Из рис. 3 видно, что для видеокарты NVIDIA Titan V доступными оказываются модели до 50 000 атомов. Для видеокарты Tesla P100 это значение несколько больше и составляет 70 000 атомов. На видеокарте Tesla A100 можно изучать модели из более чем 310 000 атомов. При этом время расчета одного шага такой модели будет на три порядка меньше времени одного шага молекулярной динамики в пакете VASP для модели  $C_{98}Li_3$ .

Таблица 2. Оценка точности DP-потенциалов на валидационных и тестовых наборах данных

Класс систем		DP-e2		DP-hybrid		DP-e3		GAP	
Углеродная поверхность	$n_{Li}$	RMSE энергии, meV/atom	RMSE силы, meV/Å	RMSE энергии, meV/atom	RMSE силы, meV/Å	RMSE энергии, meV/atom	RMSE силы, meV/Å	RMSE энергии, meV/atom	RMSE силы, meV/Å
УНТ(5,5)	1 (внутри трубки)	0.32	49	0.28	50	0.32	83	1.07	55
УНТ(5,5)	1 (вне трубки)	0.43	68	0.42	71	0.77	114	2.63	74
УНТ(5,5)	5	1.69	101	1.63	99	2.23	137	2.10	109
УНТ(5,5)	10	1.50	111	1.49	109	2.07	147	5.59	120
Графен	1	2.37	65	2.08	59	2.96	82	12.50	49
Графен	3	3.36	64	2.57	60	2.08	85	8.58	54
Бифенилен	3	0.86	93	0.77	84	0.80	133	6.75	91
Ирида-графен	3	1.05	97	0.64	76	1.11	117	9.08	82
$\psi$ -графен	3	1.04	79	0.86	72	1.29	122	5.72	75
Тетрапента-декагексагональный графен	4	0.89	87	0.73	74	1.58	140	2.44	87
Тестовый набор 1	1	4.75	95	3.14	89	5.38	156	3.12	60
Тестовый набор 2	1	3.35	20	3.31	18	3.63	21	12.67	14

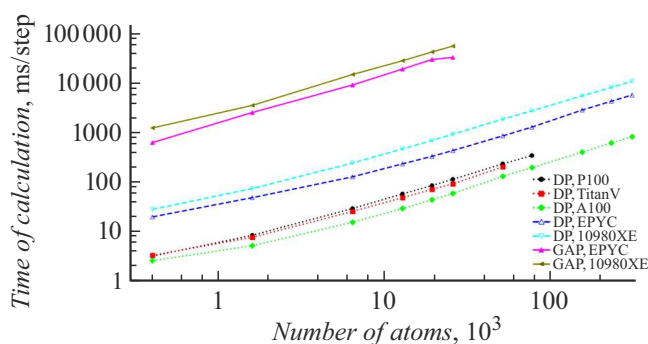


Рис. 3. Время выполнения одного шага молекулярной динамики при использовании различного аппаратного обеспечения (логарифмический масштаб осей).

Обобщая сделанные выше оценки точности полученных потенциалов (табл. 2) и требуемых ими вычислительных ресурсов (рис. 3), можно прийти к выводу, что использование потенциала GAP неоправданно даже в том случае, когда нет возможности проводить расчеты с

потенциалом DP на графических ускорителях. Если же такая возможность есть, то нецелесообразно проводить расчеты с потенциалом DP на центральном процессоре.

### 2.3. Прикладная программа

Помимо использования потенциалов DP в программе LAMMPS, они достаточно просто могут быть встроены в пользовательские программы на языке Python посредством модуля deermd. На основе ранее созданной автором программы GUI4dft [30] была разработана графическая оболочка для вычисления энергий систем с потенциалами DP, в том числе полученным в настоящей работе. Программа находится в открытом доступе на платформе GitVerse (<https://gitverse.ru/sozykin/ml-potential-calculator>). Она может быть использована для оценки энергий систем, заданных в формате VASP POSCAR, и количественной оценки степени отличия систем с помощью дескриптора SOAP. Последняя возможность доступна посредством модуля quippy.

Графический интерфейс программы приведен на рис. 4. На вкладке „Settings“ пользователь указывает

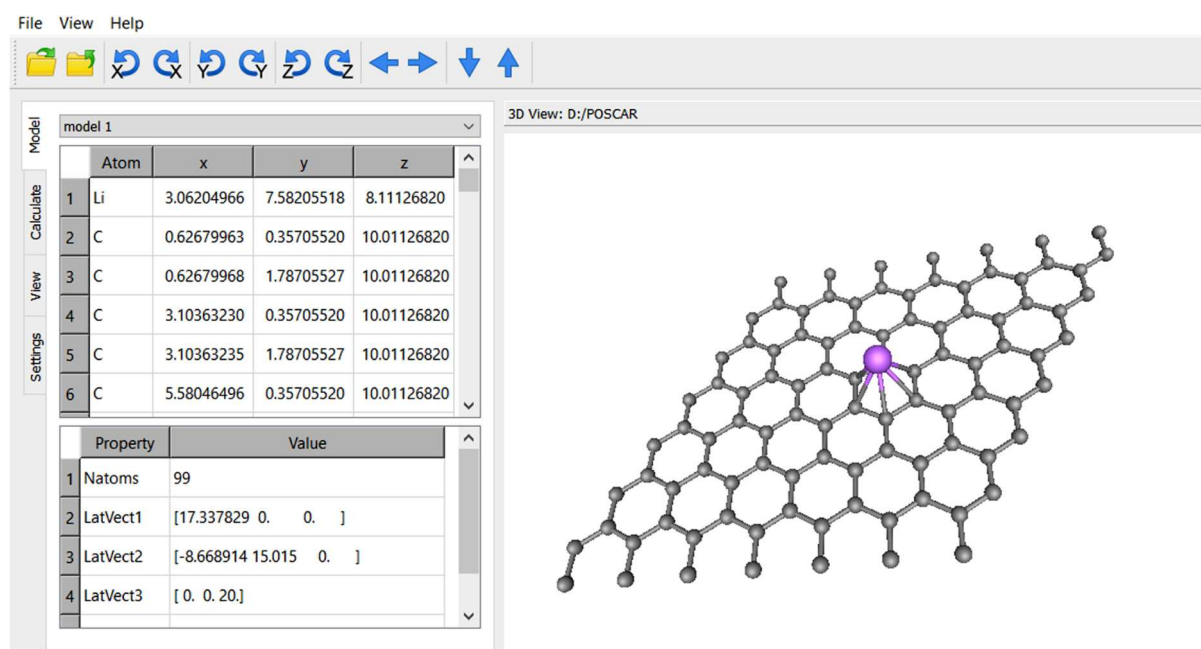


Рис. 4. Интерфейс прикладной программы.

путь к используемому потенциалу DP. На вкладке „Model“ расположена кнопка для выбора файла с атомной структурой в формате POSCAR и формы, в которых отображается информация о модели, с которой ведется работа в данное время. На вкладке „Calculation“ расположены инструменты для расчета энергий модели или дескриптора SOAP. Поскольку значение дескриптора SOAP представляет из себя вектор из большого числа элементов, его значения сложно воспринимать. Имеющиеся в программе инструменты предоставляют возможность рассчитать норму разности дескрипторов ранее открытых в программе систем. Этот параметр можно использовать в качестве меры различия систем.

Рассмотрим пример использования программы. Сравним энергии адсорбции лития, рассчитанные разными методами. Для всех неэквивалентных центров адсорбции на исследованных системах были рассчитаны энергии адсорбции лития в рамках DFT и с потенциалом DP по формуле

$$E_{ads} = E_{model+Li} - E_{model} - E_{Li},$$

где  $E_{model+Li}$  — энергия расчетной ячейки нанотрубки, графена, бифенилена, ирида-графена,  $\psi$ -графена или тетра-пента-дека-гексагонального графена с одним атомом лития,  $E_{model}$  — энергия расчетной ячейки такой же модели без атома лития,  $E_{Li}$  — энергия атома лития. Равновесные положения атома лития расположены над центрами углеродных колец поверхности. Приведем сначала результаты DFT. Для всех исследованных систем поверхности содержат шестичленные циклы, но энергии адсорбции на них отличаются: они равны  $-1.62$ ,  $-1.76$ ,

$-2.19$ ,  $-2.27$ ,  $-2.49$  и  $-2.50$  eV для графена, УНТ, бифенилена, ирида-, тетра-пента-дека-гексагонального и  $\psi$ -графена соответственно. Полученные энергии адсорбции слабо зависят от размера углеродного кольца за исключением случая адсорбции на трехчленном цикле ирида-графена, где энергия адсорбции составляет  $-1.97$  eV. В порядке увеличения размера углеродного кольца, над которым адсорбируется атом лития: бифенилен ( $-2.10$ ,  $-2.19$ ,  $-2.11$  eV), иридаграфен ( $-1.97$ ,  $-2.27$ ,  $-2.29$  eV),  $\psi$ -графен ( $-2.52$ ,  $-2.50$ ,  $-2.55$  eV), тетра-пента-дека-гексагональный графен ( $-2.43$ ,  $-2.50$ ,  $-2.49$ ,  $-2.51$  eV).

Для найденных в DFT равновесных конфигураций были рассчитаны энергии с потенциалом DP и рассчитаны энергии взаимодействия. Наименьшая средняя ошибка по разным центрам адсорбции в энергии взаимодействия была получена для тетра-пента-дека-гексагонального графена (0.05 eV), максимальная — для нанотрубки (0.32 eV). Таким образом, для данного параметра ошибка не превышает 20%.

В случае использования потенциала DP RMSE энергии для двух тестовых наборов данных, описанных выше, составили 4.75 meV/atom (95 meV/Å) и 3.35 meV/atom (20 meV/Å) для наборов 1 и 2 соответственно. Для определения областей тестовых наборов, для которых была получена наибольшая ошибка, обратимся к приведенным на рис. 5 профилям энергии. Энергии отсчитывались относительно минимального значения независимо для каждой линии. При этом для тестового набора 1 абсолютные значения энергии в минимумах практически совпадают, а для тестового набора 2 данные DFT сдвинуты относительно данных DP на величину, приблизительно

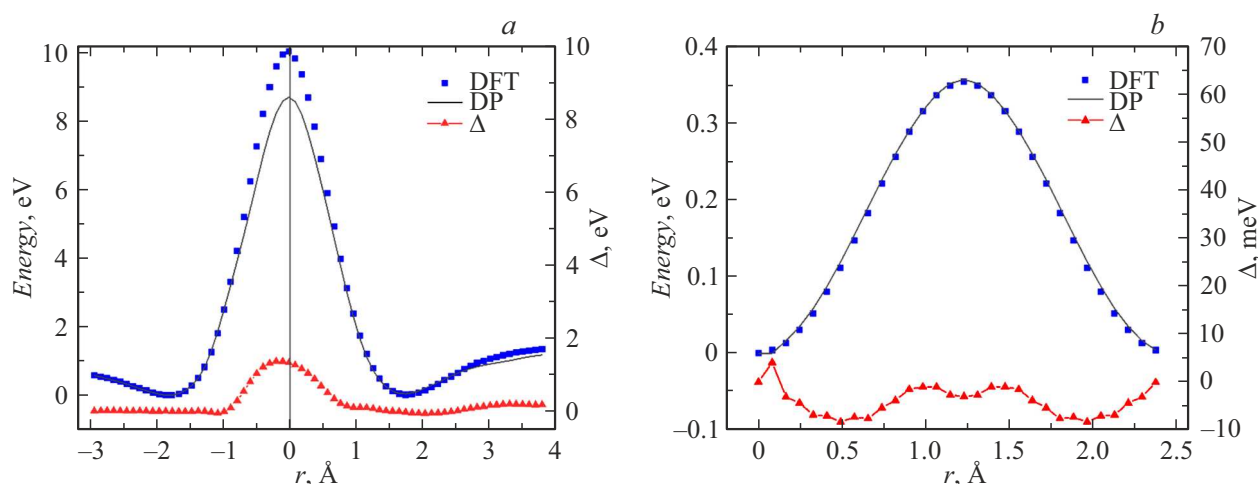


Рис. 5. DFT- и DP-энергии, их разность для тестовых наборов: *a* — 1, *b* — 2.

равную 0.33 eV. Именно этот сдвиг и отражен в RMSE. Третья линия ( $\Delta$ ) — разность относительных энергий для DFT и DP. Относительные энергии для тестового набора 2 всюду близки. Большое RMSE для первого набора связано с большими ошибками в энергии для позиций лития вблизи поверхности нанотрубки. Такие позиции далеки от тех, на которых проводилось обучение и тех, которые могут наблюдаться в эксперименте.

## Заключение

В ходе проведения исследования была создана база, содержащая данные об энергиях более чем 8700 конфигураций структур углеродных наноматериалов с атомами лития. Конфигурации возникали в ходе молекулярно-динамического моделирования при температурах до 450 K.

Сравнение двух подходов к обучению потенциалов межчастичного взаимодействия и различных дескрипторов в рамках каждого из них показало, что наиболее эффективным оказался потенциал, полученный методом глубокого обучения с двухчастичным дескриптором, учитывающим относительное положение двух атомов.

Была продемонстрирована способность с высокой точностью описывать барьеры диффузии лития на углеродной поверхности, а также энергии комплексов с литием при его расположении на типичных расстояниях от углеродной поверхности.

Сравнение энергий взаимодействия атома лития с углеродной поверхностью, рассчитанных в рамках теории функционала электронной плотности и с помощью полученного потенциала, показало, что точность определения этой величины оказывается особенно высокой (менее 0.2 eV) для случая адсорбции на графене. Максимальная ошибка (0.8 eV) была зафиксирована для случая адсорбции на иридаграфене. Указанные погрешности, однако, не превышают характерный разброс данной

величины при использовании разных первопринципных инструментов.

## Конфликт интересов

Автор заявляет, что у него нет конфликта интересов.

## Список литературы

- [1] Z. Wu, K. Sun, Z. Wang. *Batteries*, **8**, 246 (2022).
- [2] P. Lian, X. Zhu, S. Liang, Z. Li, W. Yang, H. Wang. *Electrochim. Acta*, **55** (12), 3909 (2010).
- [3] L.A. Montoro, E.Y. Matsubara, J.M. Rosolen. *J. Power Sources*, **257**, 205 (2014).
- [4] C.E. Nwanno, W. Li. *Nano Res.*, **16** (11), 12384 (2023).
- [5] M.W. Ochapski, D. Ataç, J.G.M. Sanderink, A.Y. Kovalgin, M.P. de Jong. *Carbon Trends*, **4**, 100045 (2021).
- [6] Y. Zhong, K. Deng, J. Zheng, T. Zhang, P. Liu, X. Lv, W. Tian, J. Ji. *J. Mater. Sci. Technol.*, **149**, 205 (2023).
- [7] Y. Zhou, R. Chen, Z. Gao, J. He, X. Li. *Mater. Today Energy*, **37**, 101389 (2023).
- [8] X. Yan, Z. Fu, L. Zhou, L. Hu, Y. Xia, W. Zhang, Y. Gan, J. Zhang, X. He, H. Huang. *ACS Appl. Mater. Interfaces*, **15** (14), 17986 (2023).
- [9] A.P. Bartók, M.C. Payne, R. Kondor, G. Csányi. *Phys. Rev. Lett.*, **104** (13), 136403 (2010).
- [10] A.V. Shapeev. *Multiscale Model. Simul.*, **14** (3), 1153 (2016).
- [11] J. Han, L. Zhang, R. Car, E. Weinan, *Commun. Comput. Phys.*, **23** (3), 629 (2018).
- [12] B. Deng, P. Zhong, K.J. Jun, J. Riebesell, K. Han, C.J. Bartel, G. Ceder. *Nat. Mach. Intell.*, **5** (9), 1031 (2023).
- [13] V.L. Deringer, G. Csányi. *Phys. Rev. B*, **95** (9), 094203 (2017).
- [14] P. Rowe, V.L. Deringer, P. Gasparotto, G. Csányi, A. Michaelides. *J. Chem. Phys.*, **153** (3), 034702 (2020).
- [15] T. Kocabaş, M. Keçeli, A. Vázquez-Mayagoitia, C. Sevik. *Nanoscale*, **15** (19), 8772 (2023).
- [16] H. Muhli, X. Chen, A.P. Bartók, P. Hernández-León, G. Csányi, T. Ala-Nissila, M.A. Caro. *Phys. Rev. B*, **104** (5), 054106 (2021).



- [17] Y. Wang, Z. Fan, P. Qian, T. Ala-Nissila, M.A. Caro. *Chem. Mater.*, **34** (2), 617 (2022).
- [18] J. Wang, H. Shen, R. Yang, K. Xie, C. Zhang, L. Chen, K.M. Ho, C.Z. Wang, Z. Cai, S. Wang. *Carbon*, **186**, 1 (2022).
- [19] S. Fujikake, V.L. Deringer, T.H. Lee, M. Krynski, S.R. Elliott, G. Csányi. *J. Chem. Phys.*, **148** (24), 241714 (2018).
- [20] L. Zhang, H. Wang, M.C. Muniz, A.Z. Panagiotopoulos, R. Car, E. Weinan. *J. Chem. Phys.*, **156** (12), 124107 (2022).
- [21] G. Kresse, J. Furthmuller. *Phys. Rev. B*, **54**, 11169 (1996).
- [22] I.V. Lebedeva, A.V. Lebedev, A.M. Popov, A.A. Knizhnik. *Comput. Mater. Sci.*, **128**, 45 (2017).
- [23] S. Klawohn, J.P. Darby, J.R. Kermode, G. Csányi, M.A. Caro, A.P. Bartók. *J. Chem. Phys.*, **159**, 174108 (2023).
- [24] V.L. Deringer, A.P. Bartók, N. Bernstein, D.M. Wilkins, M. Ceriotti, G. Csányi. *Chem. Rev.*, **121** (16), 10073 (2021).
- [25] Y. Shaidu, E. Küçükbenli, S. De Gironcoli. *J. Phys. Chem. C*, **122** (36), 20800 (2018).
- [26] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, E. Weinan. *Comput. Phys. Commun.*, **253**, 107206 (2020).
- [27] S. Trillat, J. Lam, S. Ispas, A.K.A. Kandy, M.E. Tuckerman, N. Tarrat, M. Benoit. *Comput. Mater. Sci.*, **236**, 112848 (2024).
- [28] M. Fronzi, R.D. Amos, R. Kobayashi, N. Matsumura, K. Watanabe, R.K. Morizawa. *Nanomaterials*, **12**, 3891 (2022).
- [29] S.K. Achar, L. Zhang, J.K. Johnson. *J. Phys. Chem. C*, **125** (27), 14874 (2021).
- [30] S.A. Sozykin. *Comput. Phys. Commun.*, **262**, 107843 (2021).