

Структурная наследственность жидкость-кристалл в потенциалах машинного обучения для сетевых систем

И. А. Балякин^{+*1)}, Р. Е. Рыльцев⁺¹⁾, Н. М. Щелкачев^{+×1)}

⁺ Институт металлургии Уральского отделения РАН, 620016 Екатеринбург, Россия

^{*} Научно-образовательный центр “Наноматериалы и нанотехнологии”,
Уральский Федеральный университет, 620002 Екатеринбург, Россия

[×] Институт физики высоких давлений им. Л. Ф. Вережагина, 142190 Москва, Россия

Поступила в редакцию 11 ноября 2022 г.

После переработки 31 января 2023 г.

Принята к публикации 31 января 2023 г.

В работе рассмотрен вопрос о способности потенциалов машинного обучения (MLIP), параметризованных с использованием только неупорядоченных конфигураций, соответствующих жидкости, описывать свойства кристаллических фаз и предсказывать их структуру. В качестве объекта исследования мы рассматриваем сетевую систему SiO₂, обладающую большим количеством полиморфных фаз, значительно отличающихся по структуре и плотности. С использованием только высокотемпературных неупорядоченных конфигураций был параметризован MLIP на основе искусственных нейронных сетей (модель DeePMD). Потенциал демонстрирует хорошую способность воспроизводить *ab initio* зависимости энергии от объема и плотности колебательных состояний для всех рассмотренных тетра- и октаэдрических кристаллических фаз SiO₂. Более того, при помощи комбинации эволюционного алгоритма и разработанного DeePMD-потенциала удалось воспроизвести реально наблюдаемые в эксперименте кристаллические структуры SiO₂. Такая хорошая переносимость жидкость-кристалл для MLIP открывает перспективы моделирования структуры и свойств новых систем, для которых отсутствует экспериментальная информация о кристаллических фазах.

DOI: 10.31857/S1234567823050099, EDN: pybaym

Введение. В последние годы достигнут существенный прогресс в атомистическом моделировании материалов, связанный с применением методов машинного обучения для построения классических межатомных потенциалов взаимодействия [1–3]. Основная идея такого подхода – аппроксимировать поверхность потенциальной энергии системы с помощью некоторой многочастичной функции общего вида (например, многослойной нейронной сети), используя эталонные значения энергий и межатомных сил, полученные с помощью *ab initio* расчетов. Данная идея аналогична предложенному в начале 1990-х методу Force-Matching [4], который в дальнейшем успешно использовался для параметризации модельных потенциалов типа погруженного атома [5–7]. Принципиальное отличие потенциалов машинного обучения (MLIP – *Machine Learning Interatomic Potentials*) в использовании гораздо более сложных многочастичных функций, имеющих иногда сотни

тысяч настраиваемых параметров, что позволяет достигать точности, сравнимой с точностью *ab initio* методов, при на порядки меньших вычислительных затратах.

Несмотря на привлекательность такого подхода, многие ключевые вопросы, касающиеся построения MLIP и их применения для решения конкретных задач, являются нерешенными. Ключевой проблемой является разработка эффективных методов формирования тренировочного датасета (ТД) – набора репрезентативных атомных конфигураций и соответствующих им значений энергий и сил, используемых для параметризации MLIP. В большинстве случаев, основой для построения ТД являются упорядоченные конфигурации, соответствующие различным кристаллическим фазам исследуемой системы; неупорядоченные (жидкие, аморфные) конфигурации используются только как дополнение к кристаллическим или в том случае, когда изучение соответствующих неупорядоченных фаз является целью исследования [3]. Возникают естественные вопросы: какие кристаллические конфигурации рассматривать

¹⁾ e-mail: i.a.balyakin@gmail.com; rrylcev@mail.ru;
n.chelchekatchev@gmail.com

для параметризации потенциалов и каким алгоритмом их генерировать. Один из возможных вариантов – использование кристаллографических баз данных, либо репозитория вычислительных проектов, таких как Materials Project, OQMD, AiiDA и др. Найденные в таких базах конфигурации формируют базовый ТД, который затем расширяется путем введения структурных дефектов и/или в ходе МД расчетов [8]. Основная проблема такого подхода – необходимость априорного знания или, по крайней мере, некоторых предположений о кристаллических структурах, которые могут возникать в исследуемой системе. Однако для многих (особенно многокомпонентных) систем подобная информация может быть весьма ограничена или вовсе недоступна. В таких случаях необходимы методы построения репрезентативных ТД, которые не используют априорную информацию о кристаллических структурах.

Один из способов решения данной задачи – параметризация MLIP с использованием только неупорядоченных конфигураций, соответствующих жидкости. В силу достаточной подвижности атомов, моделирование жидкости методом МД при различных концентрациях, температурах и давлениях позволяет достаточно быстро сформировать репрезентативный ТД и параметризовать MLIP, который позволяет с высокой точностью моделировать свойства расплавов [9, 10]. Возникает вопрос: как MLIP, параметризованный с использованием исключительно неупорядоченных жидких конфигураций, будет описывать свойства кристаллических фаз?

Данный вопрос пока мало изучен; в литературе имеется всего несколько примеров использования такого подхода. Так, недавно был предложен MLIP для моделирования рубидия, при параметризации которого использовались преимущественно неупорядоченные конфигурации [11]. Было показано, что такой потенциал воспроизводит максимум на зависимости температуры плавления ОЦК-фазы рубидия от давления. Однако рубидий – сравнительно простой пример щелочного металла с преимущественно изотропным межатомным взаимодействием. Кроме того, авторы все же использовали небольшое число кристаллических конфигураций для параметризации. Принципиальный вопрос – будет ли такая процедура машинного обучения эффективна для описания более сложных многокомпонентных систем с анизотропным взаимодействием. Частичный ответ на этот вопрос дает работа [12], где MLIP, параметризованный на конфигурациях жидкой воды, был использован для расчета плотности, энергий и колебательных спектров различных фаз льда в сравнении с *ab*

initio расчетами. Однако в данной работе наблюдается невысокая точность воспроизведения энергий кристаллических структур. А именно, на корреляционных зависимостях энергий, вычисленных с помощью MLIP и *ab initio*, имеется систематическое смещение от диагонали, равное по порядку величины дисперсии (ср. с рис. 4 в данной работе). Аналогичная проблема наблюдалась и в работе [13], где MLIP, параметризованный на жидких конфигурациях, использовался для поиска стабильных кристаллических структур с помощью эволюционного алгоритма. Это является серьезным недостатком, поскольку систематические ошибки в описании энергетики кристаллов недопустимы при использовании MLIP для предсказания их структуры.

Возникает принципиально важный вопрос: можно ли построить самосогласованный и не требующий априорной информации о кристаллических структурах метод построения MLIP, позволяющий с высокой точностью описывать как неупорядоченные, так и упорядоченные фазы и предсказывать их структуру. Очевидно, что универсального ответа на поставленный вопрос не существует и результат будет различным для систем различной природы, различных моделей машинного обучения и способов формирования ТД. В данной работе мы исследуем данный вопрос для сетевых систем, рассмотрев кремнезем SiO_2 в качестве репрезентативной системы для данного класса. Это достаточно сложная система с большим количеством полиморфных фаз, значительно отличающихся по структуре и плотности. В литературе имеются примеры построения MLIPs для SiO_2 [10, 14, 15], которые хорошо воспроизводят известные экспериментальные и расчетные данные для свойств как кристаллических, так и неупорядоченных фаз. Однако во всех случаях для моделирования кристаллов были использованы потенциалы, параметризованные с использованием соответствующих кристаллических конфигураций. Ни в одной из представленных работ не были рассмотрены вопросы о переносимости жидкость-кристалл при построении MLIPs и, в частности, о возможности предсказания структуры кристаллических фаз на основе информации о локальной структуре жидкости.

Помимо практической важности для вычислительного материаловедения, рассматриваемые вопросы имеют и чисто фундаментальный интерес для физики неупорядоченных систем. Действительно, вопрос о том, насколько локальная структура жидкостей содержит информацию о структуре кристаллов и в какой мере она может быть использована для вычисления и предсказания их свойств, яв-

ляется крайне интересным и малоизученным. В последние годы стали появляться работы, посвященные изучению данного вопроса. Оказалось, что анализ свойств расплава в некоторых случаях позволяет делать важные (в том числе и предсказательные) выводы о структуре и свойствах твердых фаз [16–19]. Для характеристики генетической взаимосвязи между структурой жидкости и кристалла мы в дальнейшем будем употреблять термин “структурная наследственность”.

Разработка потенциала. Ключевыми этапами разработки любого MLIP являются выбор способа преобразования координат локального окружения атомов в набор так называемых дескрипторов – чисел, инвариантных относительно операций трансляции, вращения и пермутации и выбор регрессионной модели, используемой для аппроксимации поверхности потенциальной энергии. Как показывает опыт, выбор дескрипторов влияет на точность и производительность модели существенней, чем выбор регрессора [9]. В данной работе мы применяем модель DeePMD [20], в которой используются две связанные искусственные нейронные сети (ИНС): сеть погружения для преобразования локального окружения атомов в дескрипторы, и аппроксимирующая сеть для преобразования дескрипторов всех атомов системы в потенциальную энергию. Архитектура обеих ИНС выбирается в виде многослойной полносвязной нейронной сети прямого распространения (так называемый перцептрон). Полное математическое описание структуры модели довольно громоздко и подробно изложено в работе [2]. Мы ограничимся только кратким описанием основных идей.

Локальное окружение i -го атома описывается матрицей $\mathcal{R}^i \in \mathbb{R}^{N_i \times 3}$, содержащей координаты N_i ближайших соседей в сфере радиуса r_{cut} , заданных в локальной системе координат. Эта матрица преобразуется в $\tilde{\mathcal{R}}^i \in \mathbb{R}^{N_i \times 4}$ как $\{\mathbf{r}_{ij}\} \rightarrow s(r_{ij})\{1, \mathbf{r}_{ij}/r_{ij}\}$, где $s(r_{ij})$ – некоторая непрерывная дифференцируемая функция, монотонно убывающая до нуля при $r \rightarrow r_{\text{cut}}$. Набор дескрипторов для окружения i -го атома формируется как $\mathcal{D}_i = \frac{1}{N_i^2} (\mathcal{G}_i^{<M})^T \tilde{\mathcal{R}}_i \tilde{\mathcal{R}}_i^T \mathcal{G}_i$, где $\mathcal{G}_i$ – вектор длины M , возвращаемый ИНС погружения, на входной слой которой подаются векторы $s(R_{ij})$, а $\mathcal{G}_i^{<M}$ – вектор, состоящий из первых $M_1 < M$ элементов вектора $\mathcal{G}_i$. Можно показать, что $\mathcal{D}_i$ обеспечивают инвариантность относительно трансляций, вращений и перестановок. Важно, что данные дескрипторы адаптивны, поскольку веса ИНС погружения также модифицируются в ходе параметризации. Для использования таких дескрипторов требуется минимум человеческого вме-

шательства, фактически необходимо лишь задание архитектуры ИНС погружения. Большое в сравнении с другими дескрипторами число параметров позволяет ИНС-дескрипторам быть достаточно гибкими для описания различных фаз материала в широком диапазоне температур и давлений.

Ключевую роль при параметризации MLIP играет способ формирования ТД. Один из стандартных способов – строить ТД из конфигураций, полученных в ходе моделирования первопринципной МД. Такой подход может давать неудовлетворительные результаты, поскольку длина траекторий в первопринципной МД, как правило, не превышает несколько пикосекунд. Генерируемые таким образом конфигурации оказываются сильно коррелированы, поскольку на временах моделирования система может оказаться запертой в области конфигурационного пространства, соответствующей некоторому локальному минимуму потенциальной энергии (вероятно, с такой проблемой столкнулись авторы работы [13]). Для решения данной проблемы применяются различные алгоритмы так называемого активного обучения, позволяющие эффективно отбирать конфигурации в широкой области конфигурационного пространства [2, 21].

В данной работе для формирования ТД использовался один из вариантов активного обучения, реализованный в программном пакете DPGEN [22]. В основе алгоритма лежит итеративное повторение трех шагов *training*, *exploration* и *labeling*. На этапе *training* на существующем ТД параметризуются несколько (N_{models}) DeePMD-потенциалов с одинаковой архитектурой ИНС, но различными случайными распределениями начальных значений параметров. Данные модели используются для проведения классического МД моделирования при различных температурах T , давлениях P и концентрациях (этап *exploration*). Использование классического МД позволяет запускать десятки параллельных расчетов при различных термодинамических параметрах и рассматривать траектории длиной в сотни пикосекунд и тем самым эффективно охватывать широкие области конфигурационного пространства. Для каждой из таких траекторий происходит контроль точности потенциала и отбор конфигураций, на которых модель дает низкую точность. Для этой цели для конфигураций $\mathcal{R}_t$, возникающих в ходе данной траектории на шаге t , оценивается величина:

$$\epsilon_t = \max_i \sqrt{|\langle F_{w,i}(\mathcal{R}_t) - \langle F_{w,i}(\mathcal{R}_t) \rangle|^2},$$

где $F_{w,i}(\mathcal{R}_t)$ – сила, действующая в DeePMD-модели с параметрами w на i атом в конфигурации $\mathcal{R}_t$,

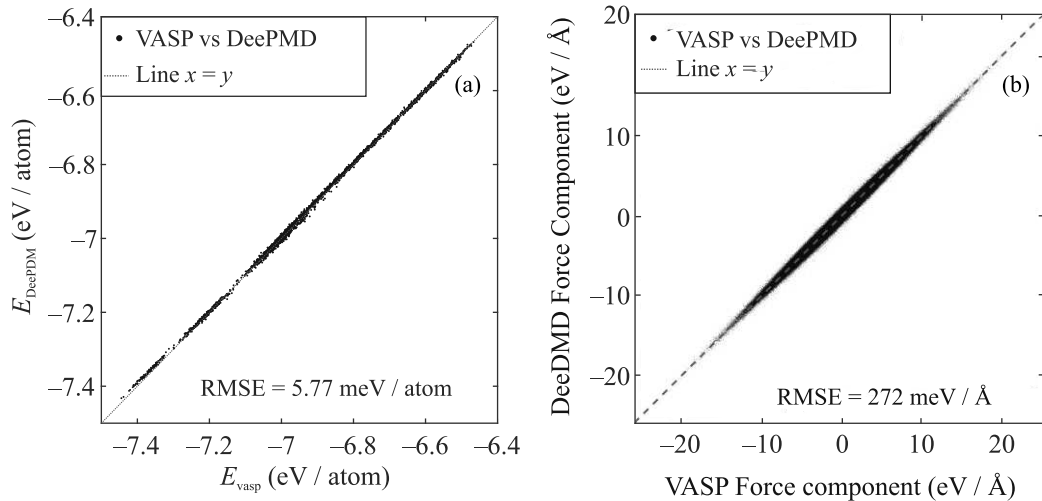


Рис. 1. (Цветной онлайн) Корреляционные зависимости энергий (a) и компонент сил (b), вычисленных в VASP и DeePMD

усреднение проводится по ансамблю моделей. Конфигурации, для которых выполняется неравенство $\epsilon_l < \epsilon_t < \epsilon_h$, отбираются в список кандидатов на пополнение ТД. Затем, для некоторого заданного числа конфигураций из списка кандидатов проводятся *ab initio* расчеты энергии, межатомных сил и тензоров напряжений (этап *labeling*). Описанная процедура повторяется, пока не наступает насыщение ТД, т.е. новые конфигурации перестают отбираться на этапе *exploration*. Данный алгоритм позволяет сформировать компактный и репрезентативный ТД, содержащий статистически независимые конфигурации.

В данной работе для проведения классического МД моделирования с MLIP использовался программный пакет LAMMPS [23]; первопринципные расчеты осуществлялись в рамках теории функционала электронной плотности (DFT), реализованной в пакете VASP [24]. В ходе параметризации рассматривались сверхъядейки, состоящие из 342 частиц (114 структурных единиц SiO_2).

Важными параметрами DeePMD модели являются: радиус обрезки r_{cut} , радиус сглаживания r_{smth} , архитектура ИНС погружения, архитектура аппроксимирующей ИНС, начальная скорость обучения l_{start} , конечная скорость обучения l_{stop} , начальный (конечный) вес энергий (сил, вириалов) $p_{e(f,v)}^{\text{start(limit)}}$, количество эпох обучения N_{epoch} , размер пакета обучения N_{batch} .

Помимо указанных выше параметров, на ИНС-потенциал влияют параметры *ab initio* расчетов. К важнейшим из них в нашем случае относятся: тип используемого приближения для обменно-корреляционной энергии, вид псевдопотенциалов,

энергия обрезки плосковолнового базиса E_{cut} , способ размытия плотности электронных состояний (ISMEAR) и параметр размытия σ , сетка в k пространстве. При расчетах в VASP использовались PAW-PBE псевдопотенциалы [25].

В таблице 1 приведены значения всех ключевых параметров активного обучения, DeePMD модели и *ab initio* расчетов. С использованием приведенных параметров было разработано два потенциала: один параметризовался только на конфигурациях расплава, соответствующих траекториям при нулевом давлении (LP-DeePMD), а второй – с включением высоких давлений вплоть до 70 ГПа (HP-DeePMD).

Таблица 1. Параметры DeePMD и VASP, используемые при параметризации потенциалов

Параметр	Значение	Параметр	Значение
T , 10^3 К	3.0–4.5	p , ГПа	0–70
ϵ_l , эВ/Å	0.15	ϵ_h , эВ/Å	0.30
r_{cut} , Å	7	r_{smth} , Å	4
ИНС погружения	[50 100]	Атомная ИНС	[250 250 250]
l_{start}	10^{-3}	l_{stop}	10^{-5}
N_{models}	4	p_e	5
p_f	500	p_v	5
N_{epoch}	10^6	N_{batch}	1
E_{cut} , эВ	520	k -сетка	Г-точка
ISMEAR	0	σ	0.05

Тестирование потенциала. На рисунке 1 представлены корреляционные зависимости энергий (рис. 1a) и компонент сил (рис. 1b), вычисленных посредством DFT и HP-DeePMD для одних и тех же конфигураций. Из рисунка 1 видно, что *ab initio*

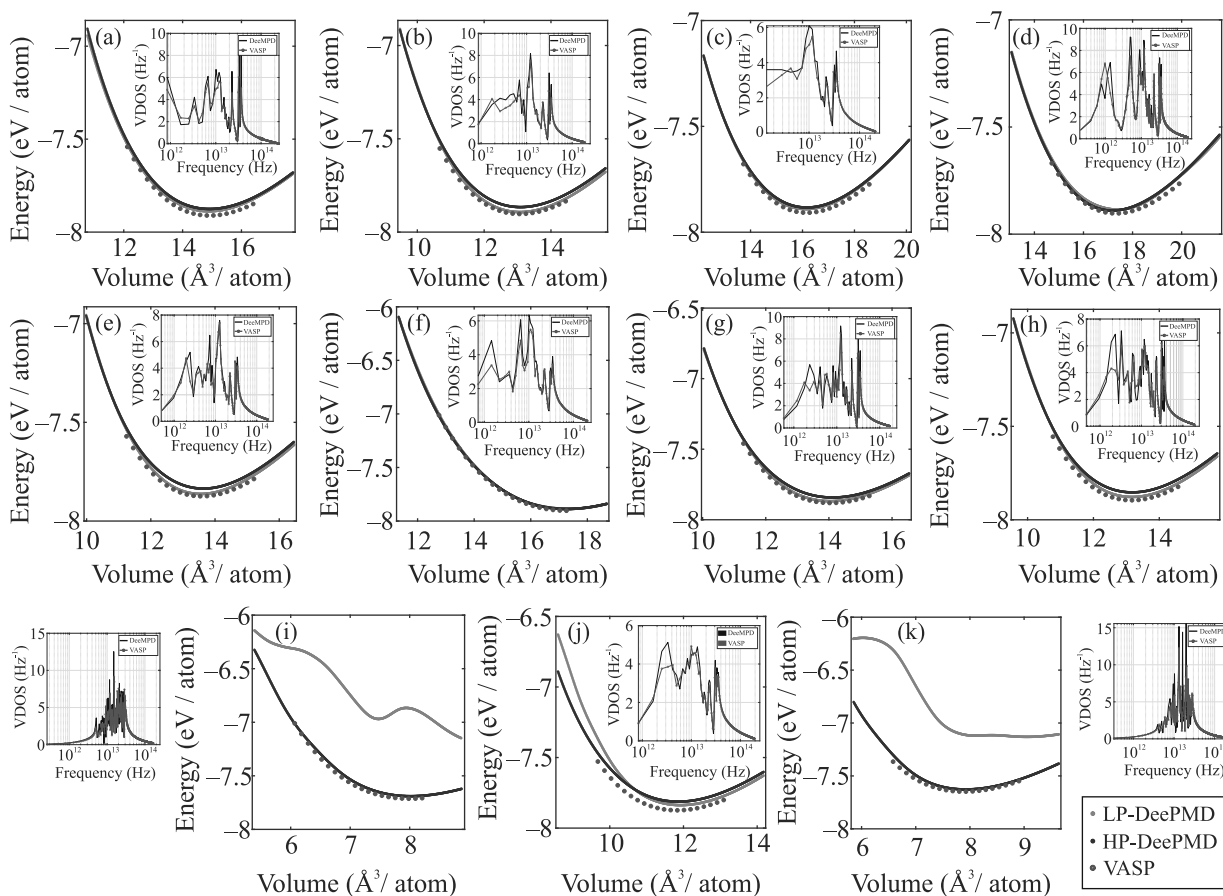


Рис. 2. (Цветной онлайн) Зависимости энергии от объема, полученные потенциалами DeePMD и DFT расчетами в VASP для: (a) – α -кварцита; (b) – α -кварца; (c) – α -тридимита; (d) – β -кварцита; (e) – β -кварца; (f) – β -тридимита; (g) – кеатита; (h) – могонита; (i) – сейферита; (j) – коэсита; (k) – стипшовита. На вставках – плотности колебательных состояний, полученные потенциалом HP-DeePMD и с помощью *ab initio* молекулярной динамики в VASP

энергии и силы достаточно хорошо воспроизводятся DeePMD-потенциалом для конфигураций, соответствующих жидкому состоянию. Опыт показывает, что такая точность обеспечивает хорошую воспроизводимость структурных и динамических свойств расплавов [9, 10].

Для анализа переносимости полученного потенциала на кристаллические фазы SiO_2 были рассчитаны зависимости энергии от объема для следующих структур: α и β кварцита, кварцы и тридимиты, кеатит, коэсит, могонит, сейферит и стипшовит. Данные зависимости, вычисленные посредством LP-DeePMD и HP-DeePMD, представлены на рис. 2. Видно, что LP-DeePMD потенциал очень хорошо воспроизводит уравнения состояния тетраэдрических фаз (рис. 2a–h, k), однако не способен описывать октаэдрические фазы (рис. 2i, k). Этот результат является ожидаемым, поскольку при нулевом давлении локальная структура расплава име-

ет только тетраэдрические конфигурации. Однако включение жидких конфигураций при высоких давлениях, которые уже содержат и октаэдрические конфигурации, позволяет описать уравнения состояния сейферита и стипшовита при практически неизменной точности описания тетраэдрических фаз. Помимо уравнений состояния, мы рассчитали плотности колебательных состояний (фурье-образы автокорреляционной функции скорости) при температурах и давлениях, находящимся в области стабильности той или иной фазы, с помощью HP-DeePMD и DFT для всех исследованных в работе кристаллических фаз (см. вставки на рис. 2). Из рисунка 2 видно, что разработанный потенциал обеспечивает хорошую точность вычисления колебательных спектров в сравнении с *ab initio* расчетами. Различия в области малых частот обусловлены относительно малой длиной траекторий *ab initio* молекулярной динамики.

Таким образом, DeePMD потенциал, параметризованный только при помощи неупорядоченных конфигураций, способен правильно описывать энергетику и колебательные свойства кристаллических фаз, что является нетривиальным результатом.

Эволюционный поиск. Для проверки возможности использования жидкостного MLIP для предсказания стабильных кристаллических структур мы использовали потенциал HP-DeePMD для оптимизации структур и вычисления их энтальпий в ходе эволюционного поиска в программном пакете USPEX [26–28]. Данным методом были определены как наиболее устойчивые, так и метастабильные кристаллические фазы при давлениях 0 и 10 ГПа. Отметим, что для SiO_2 данная задача является очень сложной, поскольку эта система, помимо фаз, присутствующих на равновесной фазовой диаграмме, имеет десятки метастабильных фаз, отличающихся по энергии на несколько мэВ/атом. По этой причине даже алгоритмы поиска, основанные на *ab initio* расчетах, могут испытывать трудности с поиском стабильных фаз в SiO_2 [29]. Так, например, в базах данных известных вычислительных материаловедческих проектов materialsproject.org и oqmd.org экспериментально наиболее стабильный α -кварц занимает лишь соответственно 56 и 29 места в сортированных по энергии списках возможных структур. Поэтому успешным тестом потенциала в ходе такого поиска является даже не правильное предсказание стабильных фаз, а хорошая корреляция между энергиями предсказанных структур, вычисленных с помощью DeePMD и VASP. На рисунке 3 представлена такая корреляционная зависимость для структур, найденных посредством HP-DeePMD и USPEX. Мы видим выраженную линейную корреляцию между *ab initio* и DeePMD энергиями (коэффициент корреляции Пирсона равен 0.987). Среднеквадратичное отклонение DeePMD значений энергий от DFT составляет 59 мэВ/атом. Однако, как видно из рис. 3, на корреляционной кривой присутствует заметное количество метастабильных структур, энергии которых отличаются от энергий экспериментально наблюдаемых фаз на несколько эВ/атом. Подобные нефизичные структуры всегда генерируются в ходе эволюционного поиска. Ясно, что их локальная структура может существенно отличаться от локальной структуры жидкости и надеяться на хорошее воспроизведение их энергий не приходится. Однако это никак не сказывается на эффективности поиска стабильных конфигураций, поскольку важным здесь является правильное описание энергетики фаз в некоторой сравнительно узкой энергетической об-

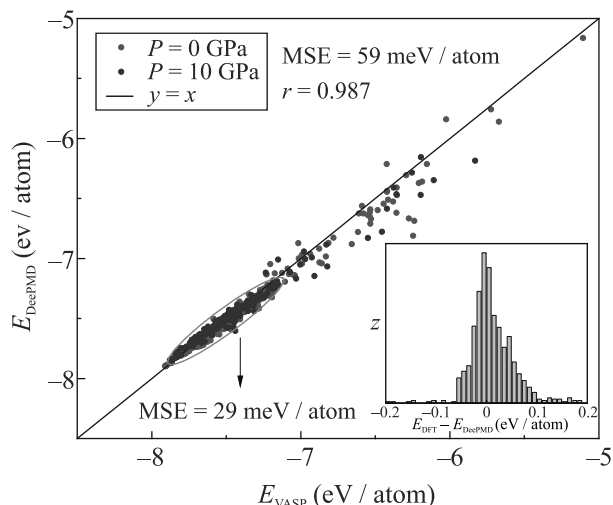


Рис. 3. (Цветной онлайн) Корреляционная зависимость DeePMD и DFT энергий кристаллических структур предсказанных USPEX с помощью DeePMD. На вставке приведена гистограмма распределения отклонения $E_{\text{DFT}} - E_{\text{DeePMD}}$. Зеленым овалом выделена область энергий, соответствующая наиболее стабильным структурам (см. диапазон энергий экспериментально наблюдаемых структур на рис. 2)

ласти. На рисунке 3 такая область отмечена зеленым овалом, ограничивающим интервал энтальпий (–8, –7) эВ/атом (см. энергии стабильных структур на рис. 2). Среднеквадратичное отклонение DeePMD значений энергий от DFT в данной области заметно меньше и составляет 29 мэВ/атом.

Анализ найденных структур позволяет заключить, что во всех случаях при $P = 0$ атомы кремния находятся в тетраэдрическом окружении атомов кислорода, а при $P = 10$ ГПа – в октаэдрическом, что наблюдается для всех известных фаз SiO_2 низкого давления и высокого давления, соответственно. Также полученные результаты воспроизводят существование десятков структур с близкими энергиями (в несколько мэВ/атом). В числе наиболее стабильных структур при нулевом давлении был найден β -кristобалит, обладающий пространственной группой симметрии 227. Наиболее стабильной структурой по результатам эволюционного поиска при 10 ГПа является структура с пространственной группой симметрии 136, что соответствует стишовиту – наиболее устойчивой фазе SiO_2 при высоких давлениях. На рисунке 4 представлены парциальные радиальные функции распределения β -кristобалита и стишовита в сравнении с соответствующими структурами из кристаллографических баз данных. Несколько смещенные в сторону больших расстояний пики расчетной структуры стишовита объясняются тем, что

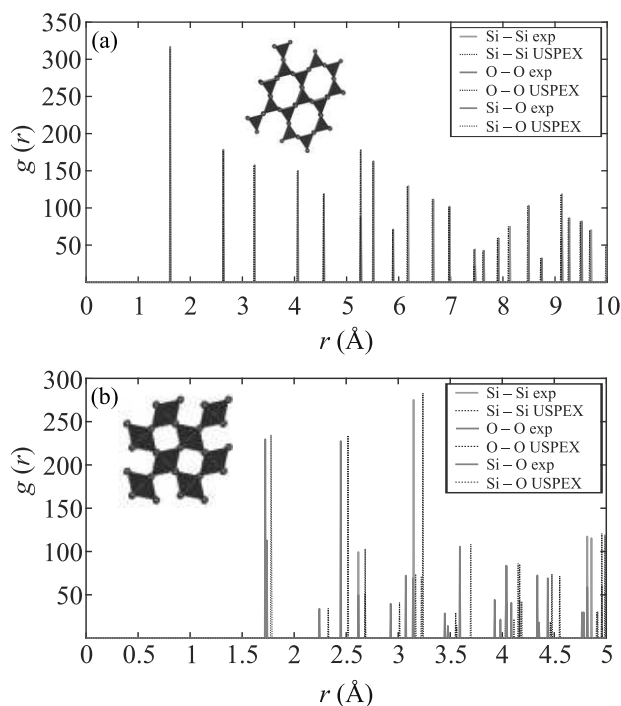


Рис. 4. (Цветной онлайн) Сравнение парциальных радиальных функций распределения для кристаллических фаз SiO_2 , найденных в ходе эволюционного поиска в программе USPEX и экспериментально наблюдаемых. (a) – β -квартц при $P = 0$ ГПа; (b) – стшовит при $P = 10$ ГПа. На вставках представлена визуализация данных структур

экспериментальный образец находился под давлением около 30 ГПа, в то время как расчет проводился для 10 ГПа.

Выводы. В данной работе изучена структурная наследственность жидкость-кристалл в потенциалах машинного обучения для сетевых систем. Показана способность MLIP, параметризованных только с помощью высокотемпературных неупорядоченных конфигураций, не только воспроизводить с *ab initio* точностью свойства кристаллических фаз, но и предсказывать их структуру при помощи эволюционных алгоритмов. А именно, с использованием DeePMD-потенциала, параметризованного по конфигурациям расплава SiO_2 в температурном диапазоне от 3000 до 4500 К и барическом диапазоне от 1 атм до 70 ГПа, удалось рассчитать с *ab initio* точностью зависимости энергии от объема для известных кристаллических фаз SiO_2 . Помимо этого, мы показали, что эволюционный поиск, реализованный в программе USPEX, позволяет предсказать структуру β -квартца и стшовита, и что даже более важно – получить хорошее описание энергетики всех найденных структур в сравнении с *ab initio* расчетами.

Полученные результаты демонстрируют глубокую связь между структурами кристалла и жидкости для сетевых систем. Такая связь, помимо фундаментального смысла, дает возможность использования MLIP, обученных на жидкости, в качестве исходной точки для построения MLIP для кристаллов. При помощи MLIP, параметризованного только на жидкости, генетическим алгоритмом может быть получен набор кристаллических структур, которые в дальнейшем должны быть включены в расширенный ТД, используемый для дообучения MLIP. Использование такой самосогласованной схемы даст возможность разрабатывать MLIP для систем, в которых априорная информация о кристаллических структурах либо отсутствует либо является ограниченной. Отметим, что для успешной практической реализации этой идеи необходимо применение алгоритмов активного обучения для создания репрезентативного набора неупорядоченных конфигураций.

Полученные результаты открывают целый ряд новых фундаментальных вопросов: как можно увеличить точность и вычислительную эффективность предложенной схемы? Будет ли данный метод работать при исследовании сложных многокомпонентных систем? Для каких классов систем структурная наследственность жидкость-кристалл достаточно ярко выражена и может быть использована для построения MLIP? Рассмотрение данных вопросов требует отдельного изучения и стимулирует развитие новых направлений исследования в вычислительном материаловедении и физике конденсированного состояния.

Исследование выполнено за счет гранта Российского научного фонда # 22-22-00506, <https://rscf.ru/project/22-22-00506/>.

Численные расчеты проводились с использованием суперкомпьютера “Уран” ИММ УрО РАН, оборудования центра коллективного пользования “Комплекс моделирования и обработки данных исследовательских установок мега-класса” НИЦ “Курчатовский институт” (<http://ckp.nrcki.ru/>) и вычислительных ресурсов Межведомственного суперкомпьютерного центра Российской академии наук (МСС РАН).

1. Y. Mishin, Acta Mater. **214**, 116980 (2021).
2. T. Wen, L. Zhang, H. Wang, E. Weinan, and D. J. Srolovitz, Materials Futures **1**, 022601 (2022).
3. V. L. Deringer, M. A. Caro, and G. Csányi, Adv. Mater. **31**, 1902765 (2019).

4. F. Ercolessi and J.B. Adams, *Europhys. Lett.* **26**, 583 (1994).
5. P. Brommer, A. Kiselev, D. Schopf, P. Beck, J. Roth, and H.R. Trebin, *Model. Simul. Mat. Sci. Eng.* **23**, 074002 (2015).
6. С.В. Стариков, В.В. Стегайлов, Г.Э. Норман et al. (Collaboration), *Письма в ЖЭТФ* **93**, 719 (2011).
7. Г.Э. Норман, С.В. Стариков, В.В. Стегайлов, *ЖЭТФ* **141**, 9102012 (2012).
8. D. Marchand, A. Jain, A. Glensk, and W.A. Curtin, *Phys. Rev. Materials* **4**, 103601 (2020).
9. R.E. Ryltsev and N.M. Chtchelkatchev, *J. Mol. Liq.* **349**, 118181 (2022).
10. I.A. Balyakin, S.V. Rempel, R.E. Ryltsev, and A.A. Rempel, *Phys. Rev. E* **102**, 052125 (2020).
11. E. Oren, D. Kartoon, and G. Makov, *J. Chem. Phys.* **157**, 014502 (2022).
12. B. Monserrat, J.G. Brandenburg, E.A. Engel, and B. Cheng, *Nat. Commun.* **11**, 1 (2020).
13. C. Hong, J.M. Choi, W. Jeong, S. Kang, S. Ju, K. Lee, J. Jung, Y. Youn, and S. Han, *Phys. Rev. B* **102**, 224104 (2020).
14. W. Li and Y. Ando, *Phys. Chem. Chem. Phys.* **20**, 30006 (2018).
15. L.C. Erhard, J. Rohrer, K. Albe, and V.L. Deringer, *npj Computational Materials* **8**, 1 (2022).
16. J. You, C. Wang, S.L. Shang, Y. Gao, H. Ju, H. Ning, Y. Wang, H.-Y. Wang, and Z.K. Liu, 10.1016/j.jma.2021.11.024.
17. R.E. Ryltsev and N.M. Chtchelkatchev, *J. Phys.: Condens. Matter* **34**, 404002 (2022).
18. V.A. Levashov, R.E. Ryltsev, and N.M. Chtchelkatchev, *Physica A* **585**, 126387 (2022).
19. L.V. Kamaeva, R.E. Ryltsev, V.I. Lad'yanov, and N.M. Chtchelkatchev, *J. Mol. Liq.* **299**, 112207 (2020).
20. H. Wang, L. Zhang, and J. Han, *Comput. Phys. Commun.* **228**, 178 (2018).
21. E.V. Podryabinki and A.V. Shapeev, *Comput. Mater. Sci.* **140**, 171 (2017).
22. Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and E. Weinan, *Comput. Phys. Commun.* **253**, 107206 (2020).
23. A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P.S. Crozier, P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M.J. Stevens, J. Tranchida, C. Trott, and S.J. Plimpton, *Comput. Phys. Commun.* **271**, 108171 (2022).
24. G. Kresse and J. Furthmuller, *Phys. Rev. B* **54**, 11169 (1996).
25. G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
26. A.R. Oganov and C.W. Glass, *J. Chem. Phys.* **124**, 244704 (2006).
27. A.R. Oganov, A.O. Lyakhov, and M. Valle, *Acc. Chem. Res.* **44**, 227 (2011).
28. A.O. Lyakhov, A.R. Oganov, H.T. Stokes, AND Q. Zhu, *Comput. Phys. Commun.* **184**, 1172 (2013).
29. A. Lahti, R. Östermark, and K. Kokko, *Comput. Mater. Sci.*, **210**, 111011 (2022).