**Шановні редакторе та рецензент!**

Ми щиро дякуємо вам за те, що ви знайшли час для рецензування нашої статті «Thermal Conductivity of Nanoporous Silicon: Molecular Dynamics Simulations and Machine Learning Prediction». Ми цінуємо ваше уважне читання та продумані відгуки, які дозволили підвищити чіткість викладу матеріалу. Ми ретельно розглянули всі зауваження та внесли відповідні зміни до тексту рукопису. Місця змін позначені червоним кольором у файлі «MarkedManuscript.pdf». Нижче ми надаємо детальні відповіді на кожне зауваження, навівши для зручності і самі коментарі. Ми сподіваємося, що переглянутий рукопис краще відповідає вашим очікуванням і стандартам для публікації в журналі «Фізика низьких температур».

**Зауваження 1.**

**1. Питання до фізичної постановки**

* **Як виникає пора?** У статті немає пояснення, чи це фізичне утворення (як результат термодинамічних процесів) чи штучно вирізана порожнина в моделі. Якщо її задають вручну, то закономірність “пора завжди посередині” — це просто умова симуляції, а не фізичний ефект. У реальних матеріалах пори розташовуються хаотично, розміри можуть варіюватися, і це залежить від температури, технології та хімічного складу.

**Відповідь 1.**

Дякуємо Рецензенту за важливе питання. У нашій моделі пора формується штучно – як сферична область всередині кристалічної решітки кремнію, з якої видалено всі атоми. Такий підхід є типовим для молекулярно-динамічних досліджень поруватих матеріалів, оскільки він дозволяє контрольовано змінювати розміри пори та ступінь пористості й оцінювати їхній вплив на теплопровідність. Розміщення пори в центрі моделі здійснюється для забезпечення максимальної симетрії та уникнення впливу меж періодичних граничних умов, що могло б спотворювати результати. Таким чином, «пора завжди посередині» є технічною умовою симуляції, яка гарантує коректність порівняння систем з різною пористістю.

Звісно, у реальних матеріалах пори можуть мати різні розміри та розташовуватись хаотично, залежно від технології синтезу та складу. Однак зазначимо, що сучасні методики дозволяють створювати й матеріали з періодичним розміщенням пор (наприклад, Journal of Power Sources Vol. 507, 2021, 230298; NPG Asia Materials Vol. 11, 2019, Article number: 12). В цій роботі ми зосереджуємося саме на фундаментальному впливі наявності пористості на теплопровідність, оскільки моделювання випадкового розташування та розподілу пор вимагає значно більших обчислювальних ресурсів і виходить за межі нашої статті, але є перспективним напрямком подальших досліджень. Відповідне доповнення додано в текст роботи у розділі Methods (перший параграф на ст.3, «*The pore was introduced as a spherical cavity by removing atoms…*»).

**Зауваження 2.**

* **Вплив радіуса пори** (тобто відношення пустих атомів до заповнених у комірці) на теплопровідність очевидний: чим більша пора - тим сильніше розсіюються фонони, тим нижча теплопровідність. Але ця залежність у реальності нелінійна та залежить від геометрії, а не тільки від об’єму порожнини.

**Відповідь 2.**

Ми погоджуємося, що у реальних зразках залежність теплопровідності від розміру та форми пор є більш складною й нелінійною. Геометрія пори відіграє суттєву роль, оскільки впливає на спектр і темп розсіяння фононів. У даній роботі ми свідомо розглядаємо спрощену модель із сферичною порою в центрі кристалу, змінюючи її радіус. Такий підхід дозволяє виокремити базовий тренд зниження теплопровідності зі збільшенням пористості та кількісно оцінити силу цього ефекту в умовах максимальної симетрії. Дійсно, нелінійність проявляється навіть у такій ідеалізованій системі, однак у нашій моделі ми отримали монотонне зменшення теплопровідності зі зростанням радіуса пори.

**Зауваження 3.**

**2. Чому передбачати всю криву k(tc)**

* Фінальна мета — отримати граничне значення інтегралу (плато k при великому tc ​). Можна було б будувати модель, яка одразу передбачає це плато, без реконструкції всієї кривої. Поточний підхід виглядає як зайвий обхідний шлях, хоча, можливо, вони так робили для кращого контролю форми кривої.

**Відповідь 3.**

Щиро вдячні Рецензенту за увагу до цього аспекту роботи. Справді, метою є отримання граничного значення інтегралу, яке відповідає плато коефіцієнта теплопровідності при великих значеннях часу кореляції *tc*. Зауважимо, що у методі Ґріна–Кубо побудова повної інтегральної кривої є загальноприйнятою процедурою, оскільки вона дозволяє не лише визначити асимптотичне значення *k*, але й перевірити стабільність результату, а також наявність шуму чи коливань на довших часових масштабах. Вказаний підхід забезпечує контроль якості розрахунків і дозволяє уникнути систематичних похибок, які могли б виникати при прямому екстраполюванні плато без аналізу поведінки всієї функції.

**Зауваження 4.**

**Критичні питання:**

**3. Питання до вибору метрик**

* Автори використали MWSE (mean weighted squared error) замість стандартного MSE. Якщо ваги підібрані так, щоб приділяти більше уваги певним ділянкам кривої (наприклад, плато), то це може бути способом “підтягнути” модель під бажану форму кривої. Такий вибір треба пояснювати чітко, інакше є ризик штучного покращення метрик.

**Відповідь 4.**

Рецензент абсолютно правий, що стандартною похибкою, яка оцінюється при тренуванні / тестуванні моделей машинного навчання є MSE. Проте мусимо зауважити, що іншою стандартною практикою є нормалізація та масштабування даних (як вхідних дескрипторів, так і цільової змінної), наприклад з використанням Min-Max Scaling чи Z-score Standardization. У випадку, коли певні величини різняться на порядки, клисичним підходом є використання log-трансформації перед масштабуванням. Це дозволяє прискорити навчання, уникнути числової нестабільності та зробити внесок різних ознак (або різних значень ознаки чи цільової функції) порівнянним. Саме такий підхід і був нами використаний при застосуванні методів RF, GB та SVR для передбачення залежностей k(tc): попередня обробка як дескрипторів так і цільової змінної проводилася таким чином, щоб у кожний набір характеризувався одиничною дисперсією та нульовим середнім значенням, причому для значень часу попередньо застосовувалася log-трансформація, а при налаштування моделей використовувалося значення MSE.

Проте при використанні методу Symbolic Regression згаданий підхід до попередньої обробки даних не дуже доцільний, так як метою було отримання аналітичного виразу, який би максимально прозоро описував залежність теплопровідності від температури та поруватості. Водночас діапазон цільових значень теплопровідності ТС фактично охоплював 2 порядки (від 1 до 226 Вт/мК). Крім того, абсолютні похибки розрахованих за допомогою MD значень, також зростали зі збільшенням величини теплопровідності – див. рис.3. Тому для того, щоб запобігти пристосуванню моделі насамперед до великих значень ТС (малих значень поруватості та температури), при налаштуванні моделі Symbolic Regression використовувався зважений метод найменших квадратів, де як ваги, застосовувалися значення оберненні до величин ТС. Тобто мова не йде про пристосування до певних ділянок кривої, а лише про врахування широкого діапазону змін цільової змінної. Зазначимо, що використання вагових коефіцієнтів є стандартною можливість пакету PySR і є одним із рекомендованих способів налаштування моделі (https://astroautomata.com/PySR/tuning/).

В тексті рукопису додана інформація щодо причин використання вагових коефіцієнтів (ст.4, останній абзац) та уточнено опис застосування класичних методів машинного навчання (ст.5).

**Зауваження 5.**

**4. Перевірка фізичної адекватності на простих моделях**

* Для простих випадків (наприклад, монокристал без пор) теплопровідність можна оцінити аналітично (Boltzmann Transport Equation, Debye model). При пористості → 0 модель має відтворювати теплопровідність монокристалу в межах допустимої похибки. При дуже великій пористості — теплопровідність повинна прямувати до нуля.
* Запустити навчання кілька разів із різними початковими умовами. Якщо форма кривих суттєво змінюється — модель нестабільна.
* Додати невеликий випадковий шум до входів (p, T) і перевірити, чи змінюються передбачення сильно.
* Побудувати learning curves (похибка від розміру навчальної вибірки), щоб перевірити деградацію при видаленні частини даних.

**Відповідь 5.**

Ми повністю згодні з Рецензентом щодо необхідності перевірки фізичної адекватності отриманих результатів. Зокрема, із загально фізичних міркувань та літературних даних очікується, що а) теплопровідність має зменшуватися при зростанні температури; б) залежність k(T) має модифікуватися зі збільшенням неупорядкованості; в) підвищення поруватості має викликати послаблення здатності матеріалу переносити теплову енергію. Зазначимо, що всі ці тенденції є характерними як для результатів MD розрахунків, так і для передбачень, отриманих за допомогою методів Symbolic Regression, Random Forest та Gradient Boosting. Єдине виключення пов’язане з Support Vector Regression, який пропонує явно завищені значення теплопровідності при високих температурах та ступенях поруватості. Більше того, 1) отриманий показник ступеня температурної залежності теплопровідності монокристалічного кремнію (1,15) близький до 1, яка очікується для напівпровідників (про що сказано у першому параграфі ст.8 рукопису); 2) отримані абсолютні значення теплопровідності монокристалічного кремнію не кардинально перевищують експериментальні дані (див. рис.3), причому саме перевищення є фізично виправданим через бездефектність та моноізотопність структур, для яких проводилися розрахунки на відміну від реальних кристалів. Таким чином, фізична адекватність отриманих результатів на нашу думку безсумнівна.

Щодо багатократного запуску процесу навчання, то фактично такий підхід і використовувався: у випадку символічної регресії у рукописі представлено типовий результат для 20 запусків; при налаштуванні класичних ML моделей та оцінці їхніх здатностей до передбачень використовувалась 5-кратна крос-валідація. Відповідь на питання щодо корисності використання випадкового шуму на вході, на нашу думку, не очевидне. По-перше, алгоритми Random Forest та Gradient Boosting нечутливі до перенавчання, а в цьому випадку, як відомо, зашумлення практично не допомагає. По-друге, розмір тренувальний набір для Symbolic Regression достатньо малий в тому додавання шуму у недоцільне, оскільки воно спотворює обмежену інформацію, підвищує нестабільність і ризик переускладнення моделі, не створює нової інформації, а лише розмиваючи наявні закономірності. Невеликий розмір навчального набору перешкоджає і якісному використанню learning curves, оскільки зменшення тренувальної підмножини робить оцінку похибки надто нестабільною, а сама крива втрачає інформативність і не відображає реальної динаміки навчання. Хоча зрозуміло, що збільшення навчальної вибірки значень теплопровідності було б корисним. Проте в нашому випадку існує обмеження щодо наявних розрахункових потужностей і саме для подолання цих проблем було запропоноване використання алгоритмів машинного навчання. Для випадку RF, GB та SVR знову ж таки згадаємо про використану крос-валідацію, яка саме передбачає видалення частини даних при тренуванні.

В рукописі додано інформацію щодо відповідності отриманих за допомогою моделей машинного навчання результатів фізичним очікуванням (ст.11).

**Зауваження 6.**

**5. Валідація**

* Для діапазонів T,pT, де немає MD-результатів, вони порівнюють RF/GB/SVR не з реальними даними, а з рівнянням (7) із Symbolic Regression (SR). Тобто, по суті, оцінюють, наскільки ML-моделі згодні з іншою моделлю, а не з фізичною “істиною”.
* Якщо SR помиляється в якійсь області, то всі метрики фактично вимірюють “згоду моделей між собою”, а не правильність результату.
* Використання 10 000 точок для одного набору (T, p) у тесті може штучно занижувати похибку через сильну кореляцію даних.

**Відповідь 6.**

Ми повністю погоджуємося, що порівняння класичних ML-моделей з Symbolic Regression у діапазонах, де відсутні MD-результати, фактично оцінює не узгодженість із фізичною «істинною», а узгодженість моделей між собою. Водночас слід зазначити, що як SR, так і RF/GB/SVR демонструють достатньо низькі похибки на тренувальному наборі, який ґрунтується на достовірних MD-розрахунках (для SR метрики наведені на рис.6а, тоді як для GB, RF та SVR МАРЕ складає 1.05%, 0.98% та 26.8 %, а R2 0.999, 0.999 та 0.977 відповідно). Це дозволяє вважати SR у цих діапазонах надійною аппроксимацією фізичної залежності, а отже, порівняння інших ML-моделей із SR дає змістовну оцінку їхньої здатності узагальнювати тренувальні дані. Зазначимо, що це повністю не замінює перевірку на MD-даних (або інших), проте дозволяє наочно продемонструвати поведінку та фізичну правдоподібність прогнозів моделей у широкому діапазоні параметрів, де прямі розрахунки відсутні. Як видно з отриманих даних, якщо SR помиляється в якійсь області, то використання RF та GB для побудови відповідних залежностей k(tc) дозволяє отримати альтернативний шлях оцінки теплопровідності.

Ми також погоджуємося, що часові ряди з великою кількістю точок можуть містити значну автокореляцію, і тому проста оцінка похибки за всіма точками справді може створювати враження заниженої невизначеності. Водночас у нашому випадку використання 10 000 точок відповідає природній дискретизації MD-симуляцій і необхідне для стабільного визначення інтегральної характеристики — теплопровідності. Тобто дані мають надлишковість, але саме вона забезпечує надійне чисельне інтегрування та згладжує флуктуації. Крім того, ми оцінюємо точність моделей не лише на основі кожної точки часового ряду, а й за узагальненою величиною теплопровідності, яка отримана з урахуванням усієї динаміки. Таким чином, хоча міжсімплова кореляція присутня, вона не призводить до штучного заниження похибки ключового результату.

У рукописі додано значення метрик для передбачень теплопровідності, отриманих за допомогою RF/GB/SVR, для значень температури та поруватості, для яких наявні MD-розрахунки (ст.10, другий абзац зверху), а також обгрунтування доцільності порівняння прогнозів для інших пар (Т, р) з формулою, отриманою за допомогою символьної регресії (ст.10, останній абзац).