

**ВИСНОВОК**  
про наукову новизну, теоретичне та практичне значення результатів  
дисертації Бердника Михайла Ігоровича  
на тему «Метод L<sub>1</sub>-регуляризації для опису фізико-хімічних властивостей молекул»  
на здобуття ступеня доктора філософії  
за спеціальністю 102 – Хімія  
з галузі знань 10 – Природничі науки

**1. Обґрунтування вибору теми дослідження та її зв'язок із планами наукових робіт університету.**

Актуальність теми дисертаційної роботи обумовлена тим, що сучасні дослідження фізичної хімії неможливі без побудови теоретичних моделей тих властивостей, що вивчаються. Такі моделі, на сьогоднішній день, мають надзвичайне різноманіття. Серед поширеніших теоретичних підходів фізичної хімії і статистичні дослідження за типом «структурно-активність» і методи молекулярної динаміки і квантово-хімічні розрахунки молекул та комплексів. Незважаючи на таку досить строкату картину теоретичних методів вони усі мають певну загальну характерну рису. А саме – теоретична модель завжди спирається на певний набір параметрів – незалежних змінних. Функція від цих параметрів і характеризує специфіку моделі. Так, якщо маємо справу із суто статистичним підходом де параметри це молекулярні дескриптори, які характеризують різноманітні аспекти структури, то функція це регресійне рівняння, або нейронна мережа. Якщо мова іде про квантово-хімічні розрахунки то співвідношення Релея визначає залежність енергії від хвильової функції яка певним чином параметризована в залежності від специфіки методу. При цьому кількість таких параметрів може бути надзвичайно великою в залежності від розрахункового методу.

Таким чином постає досить загальна актуальна проблема скорочення кількості параметрів таким чином, щоб: а) зберегти фізично важливі компоненти що гарантують адекватність розрахунку; б) спростити задачу в сенсі розрахункових (комп’ютерних) витрат; в) отримати новий, скорочений набір параметрів який дозволить надати спрощену інтерпретацію результатів.

Таким чином, тема дисертаційної роботи є актуальною, оскільки в ній вперше досліджено і доведено можливості використання прийому L<sub>1</sub>-регуляризації як методу скорочення незалежних параметрів у використанні до проблем хімії.

Актуальність описаних проблем, їх теоретичне та практичне значення зумовили вибір теми дослідження, його мету та завдання.

**Мета дисертаційної роботи** полягала у вивченні можливостей використання L<sub>1</sub>-регуляризації до багато параметричних задач, що виникають в хімії.

Для досягнення цієї мети в роботі поставлені та вирішенні такі **завдання**:

- програмно реалізувати основні розрахункові алгоритми L<sub>1</sub>-регуляризації;
- програмно реалізувати й дослідити різні альтернативні способи побудови лінійної регресії;
- програмно реалізувати й дослідити методи валідації отриманих регресійних рівнянь;

- розробити, програмно реалізувати й тестувати метод побудови класифікаційних функцій на основі  $L_1$ -регуляризації;
- На основі реалізованого програмного комплексу виконати розрахунки фізико-хімічних властивостей молекул серед яких є органічних систем різної природи, температури кипіння та ін;
- дослідити можливість використання  $L_1$ -регуляризації в квантовій хімії

**Об'єктом дослідження** є фізико-хімічні властивості молекул, як-от: температури кипіння, константи іонізації органічних сполук; енергії малих молекулярних систем.

**Предметом дослідження** є прогностичні моделі отримані за використання підходів  $L_1$ -регуляризації; методи побудови альтернативних рівнянь лінійної регресії; порівняльний аналіз регресійних підходів;  $L_1$ -регуляризовані багаточастинкові квантово-хімічні методи (MP2, CCSD).

**Методи дослідження:** лінійний регресійний аналіз як метод побудови моделей, що здатні прогнозувати величини активностей/властивостей; логістичний регресійний аналіз для побудови класифікаційних функцій; багаточастинкові методи квантової хімії, як-от: MP2, CCD, CCSD, які здатні адекватно описати енергетичні характеристики молекул. Квантово-хімічні методи оптимізації геометрії органічних молекул.

## 2. Формулювання наукового завдання, нове вирішення якого отримано в дисертації.

Дисертаційна робота Михайла Бердника присвячена розробці підходу, який дозволяє провести скорочення незалежних змінних, як-от: дескрипторного набору або конфігураційного розкладу хвильової функції багаточастинкового методу урахування електронної кореляції для проблем хімії.

## 3. Наукові положення, розроблені особисто дисертантом, та їх новизна.

Наукова новизна результатів дослідження, отриманих особисто здобувачем, полягає у наступному:

- Показано, що  $L_1$ -регуляризація здатна створити ранжований список молекулярних дескрипторів, завдяки чому можлива побудова простих малопараметричних регресійних рівнянь на основі методу найменших квадратів, методу найменших модулів, методу ортогональних відстаней (OLS, LAD, ODR).
- Показано, що отримані на основі алгоритму LARS-LASSO малопараметричні регресійні рівняння (OLS, LAD) можуть мати значно кращі прогностичні характеристики ніж стандартні методи, у яких не робиться відбір дескрипторів (наприклад неповний метод найменших квадратів, PLS).
- Представлено умови, за яких розділення вхідних даних на тестову й тренувальну вибірки гарантують адекватну оцінку обраної моделі.
- Показано, що алгоритм LARS-LASSO може бути з успіхом використано в побудові компактного рівняння логістичної регресії для бінарної класифікації молекул за активністю. Представлено прості рівняння, за допомогою яких можлива класифікація лігандів receptorів естрогену на активні і неактивні.
- Вперше на прикладі методів MP2, CCD, CCSD показано, що за допомогою  $L_1$ -регуляризації можливе створення впорядкованого списку електронно-

збуджених конфігурацій, який може бути використаний для створення прогресивної системи наближень квантово-хімічного методу.

#### **4. Обґрунтованість і достовірність наукових положень, висновків і рекомендацій, які захищаються.**

Обґрунтованість та достовірність наукових положень, результатів і висновків дисертації забезпечена коректним застосуванням алгоритмів запропонованих в літературі, а також співставленням одержаних теоретичних та експериментальних результатів.

#### **5. Рівень теоретичної підготовки здобувача, його особистий внесок у вирішення конкретного наукового завдання. Рівень обізнаності здобувача з результатами наукових досліджень інших учених.**

У процесі виконання наукових досліджень здобувач продемонстрував високий рівень теоретичної підготовки у галузі хемометрії, хемоінформатики, а також квантовій хімії.

Аспірантом було успішно розв'язано низку теоретичних задач в результаті чого створено відповідні комп'ютерні програми загального призначення. А саме, як результат дисертаційної роботи, базуючись на  $L_1$ -регуляризаційному підході, розроблено пакет статистичних програм що реалізують методи лінійної та логістичної регресій, та квантово-хімічних програм які здатні надати адекватні оцінки фізико-хімічних параметрів за відносно невеликих затрат комп'ютерних ресурсів.

Здобувач продемонстрував високий рівень обізнаності з науковими роботами інших учених, що використовували метод  $L_1$ -регуляризації в різних галузях знань.

#### **6. Наукове та практичне значення роботи.**

- Показано, що  $L_1$ -регуляризація здатна створити ранжований список молекулярних дескрипторів, завдяки чому можлива побудова простих малопараметричних регресійних рівнянь на основі методу найменших квадратів, методу найменших модулів, методу ортогональних відстаней (OLS, LAD, ODR).
- Показано, що отримані на основі алгоритму LARS-LASSO малопараметричні регресійні рівняння (OLS, LAD) можуть мати значно кращі прогностичні характеристики ніж стандартні методи, у яких не робиться відбір дескрипторів (наприклад неповний метод найменших квадратів, PLS).
- Представлено умови, за яких розділення вхідних даних на тестову й тренувальну вибірки гарантують адекватну оцінку обраної моделі.
- Показано, що алгоритм LARS-LASSO може бути з успіхом використано в побудові компактного рівняння логістичної регресії для бінарної класифікації молекул за активністю.
- Вперше показано, що за допомогою методу  $L_1$ -регуляризації можливе створення впорядкованого списку електронно-збуджених конфігурацій, який може бути використаний для створення прогресивної системи наближень квантово-хімічного методу (на прикладі MP2, CCD, CCSD).

## **7. Використання результатів роботи.**

Протягом виконання дисертаційної роботи результати досліджень були використані в НДР кафедри хімічної матеріалознавства Харківського національного університету імені В. Н. Каразіна: "Органічні модифікатори та іон-молекулярні системи й нові матеріали на їх основі для аналітичного та електрохімічного застосування" № держреєстрації 0118U002025.

## **8. Повнота викладу матеріалів дисертації в публікаціях та особистий внесок здобувача в публікації.**

Основні положення дисертації опубліковані у 13 наукових працях, серед яких: 4 статті у наукових фахових виданнях України (серед яких 1 у виданні, що входить до міжнародної наукометричної бази Scopus), 1 стаття у періодичному науковому виданні держави-члена ЄС, що входить до міжнародної наукометричної бази (Scopus), 8 матеріалів та тез доповідей на конференціях.

### **Публікації у науковоому фаховому виданні України, що входить міжнародної наукометричної бази Scopus:**

- [1] Berdnyk, M. I.; Zakharov, A. B.; Ivanov, V. V. Application Of L<sub>1</sub>- Regularization Approach In QSAR Problem. Linear Regression And Artificial Neural Networks. *Methods Objects Chem. Anal.* **2019**, 14 (2), 79–90. <https://doi.org/10.17721/moca.2019.79-90>.

(Особистий внесок здобувача: програмна реалізація застосованих методів регресії а також штучних нейронних мереж, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, участь у обговоренні результатів, написання публікації).

### **Наукові праці, в яких опубліковані основні наукові результати дисертації**

- [2] Бердник, М. И.; Иванов, В. В. Многошаговые Методы Первого Порядка в Решении Уравнений Теории Связанных Кластеров. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія Хімія* **2015**, № 25, 39-45.

(Особистий внесок здобувача: програмна реалізація методу зв'язаних кластерів, а також методів оптимізації до впроваджених ітеративних схем, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання публікації).

- [3] Бердник, М. И.; Иванов, В. В. L<sub>1</sub>-Регуляризация в Квантовой Химии. π-Электронная Теория Связанных Кластеров с Учетом Двукратных Возбуждений. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія Хімія* **2016**, № 26, 58–64.

(Особистий внесок здобувача: програмна реалізація L<sub>1</sub>-регуляризованого методу зв'язаних кластерів, розрахунки з використанням програмно-реалізованого методу, участь у обговоренні результатів, написання публікації).

- [4] Berdnyk, M. I.; Onizhuk, M. O.; Ivanov, V. V Methods for Building Linear Regression Equations in the “Structure-Property” Problems. *Kharkov Univ. Bull. Chem. Ser.* **2018**, № 30, 6–17. <https://doi.org/10.26565/2220-637x-2018-30-01>.

(Особистий внесок здобувача: програмна реалізація застосованих методів регресії, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання публікації).

**Наукові праці, в яких опубліковані основні наукові результати дисертації публікації у періодичних наукових виданнях інших держав, що входять до ОЕСР, і реферуються у міжнародних наукометрических базах**

- [5] Ivanov, V. V.; Berdnyk, M. I.; Adamowicz, L. L<sub>1</sub>-Regularisation of the Coupled-Cluster Solutions. *Mol. Phys.* **2017**, *115* (21–22), 2892–2902.  
<https://doi.org/10.1080/00268976.2017.1359345>.

(Особистий внесок здобувача: програмна реалізація L<sub>1</sub>-регуляризованого методу зв'язаних кластерів, розрахунки з використанням програмно-реалізованого методу, участь у обговоренні результатів, написання публікації).

**Наукові праці, які засвідчують апробацію матеріалів дисертації**

- [6] Бердник, М. И.; Иванов, В. В. L<sub>1</sub>-регуляризация. от статистики до квантовой химии, *Хімічні Каразінські читання - 2016* : тези доп. VIII всеукр. наук. конф. студентів та аспірантів, Харків, Україна, квітень 18–20, 2016; ХНУ імені В. Н. Каразіна: Харків, 2016; С. 132-133.

(Особистий внесок здобувача: програмна реалізація L<sub>1</sub>-регуляризованих статистичних методів і методу зв'язаних кластерів, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

- [7] Бердник, М. И.; Иванов, В. В. Применение l<sub>1</sub>-регуляризации в неэмпирических и полуэмпирических расчетах квантовой химии, *XII Всеукраїнська конференція молодих вчених та студентів з актуальних питань хімії* : збірка праць всеукр. наук. конф., Харків, Україна, травень 11-13, 2016; ДНУ НТК ІМК НАНУ: Харків, 2016; С. 32.

(Особистий внесок здобувача: програмна реалізація L<sub>1</sub>-регуляризованого методу зв'язаних кластерів, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

- [8] Бердник, М.И.; Дяченко, А.В.; Иванов, В. В; Регрессионные модели QSAR, *Збірник тез доповідей, Хімічні Проблеми Сьогодення (ХПС-2018)*, Вінниця, Україна, березень 27-29, 2018; Донецький національний університет імені Василя Стуса: Вінниця, 2018; С. 177.

(Особистий внесок здобувача: програмна реалізація методу LARS-LASSO а також методів регресії, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

- [9] Berdnyk, M.; Ivanov, V.; Zakharov, A; L<sub>1</sub>-Regularization In Different Applications Of Chemical Modeling, *Molecular Engineering And Computational Modelling For Nano- And Biotechnology: From Nanoelectronics To Biopolymers* : Book of Abstracts International Scientific Conference, Cherkasy, Ukraine, September 25–26, 2018; Bohdan Khmelnytsky Cherkasy National University: Cherkasy, 2018; P. 30-33.

(Особистий внесок здобувача: програмна реалізація методів регресії, регуляризованих квантово-хімічних методів, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

- [10] Бердник, М.И.; L<sub>1</sub>-регуляційний підхід у розрахунках фізикохімічних властивостей молекул, *Сучасні Проблеми Хімії* : тези доповідей XX Міжнародної конференції

студентів та аспірантів, Київ, Україна, травень 15–17 травня, 2019; Київський національний університет імені Тараса Шевченка: Київ, 2019; С. 140.

(Особистий внесок здобувача: програмна реалізація використаних методів, розрахунки фізико-хімічних властивостей молекул, участь у обговоренні результатів, написання тез, доповідь на конференції).

[11] Berdnyk, M. I.; Denysenko, K. A.; Zakharov, A. B.; Ivanov, V. V.; Validation Of Regression Equations In QSAR Problem, *Сучасні Тенденції 2020* : Тези доповідей Київської Конференції з аналітичної хімії, Київ, Україна, жовтень 21-23, 2020; Київський національний університет імені Тараса Шевченка: Київ, 2020; С.79-80.

(Особистий внесок здобувача: програмна реалізація методів валідації, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

[12] Денисенко, К. А.; Бердник, М. И.; Захаров, А. Б.; Метод валидации уравнений линейной регрессии, *Хімічні Каразінські читання - 2021* : тези доп. XIII всеукр. наук. конф. студентів та аспірантів, Харків, Україна, квітень 20–21, 2021; ХНУ імені В. Н. Каразіна: Харків, 2021; С. 122-123.

(Особистий внесок здобувача: програмна реалізація методів валідації, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, участь у написанні тез).

[13] Berdnyk, M.; Ivanov, V.; Application Of Lasso Logistic Regression To Classification Problems In Chemistry, *Modern Chemistry Problems* : Book of abstracts XXII International Conference for Students, PhD Students and Young Scientists, Київ, Україна, травень 19–21, 2021; Київський національний університет імені Тараса Шевченка: Київ, 2021; С. 9.

(Особистий внесок здобувача: програмна реалізація методів класифікації, розрахунки з використанням програмно-реалізованих методів, участь у обговоренні результатів, написання тез, доповідь на конференції).

Результати дисертаційної роботи повністю відображені в наукових працях здобувача.

**На підставі вивчення тексту дисертації здобувача, наукових праць здобувача та Протоколу контролю оригінальності (перевірку наявності текстових запозичень виконано в антиплагіатній інтернет-системі StrikePlagiarism) встановлено, що дисертаційна робота виконана самостійно, текст дисертації не містить плагіату, а дисертація відповідає вимогам академічної добросесності.**

## **9. Апробація матеріалів дисертації.**

Основні результати роботи були представлені на VIII Всеукраїнській науковій конференції студентів та аспірантів «Хімічні Каразінські читання – 2016» (Харків, 2016), XII Всеукраїнській конференції молодих вчених та студентів з актуальних питань хімії (Харків, 2016), I Міжнародній (XI Української) науковій конференції студентів, аспірантів і молодих вчених «Хімічні проблеми сьогодення» (Вінниця, 2018), міжнародній науковій конференції: «Молекулярна інженерія та комп'ютерне моделювання для нано- і біотехнологій: від наноелектроніки до біополімерів» (Черкаси, 2018), XX Міжнародній конференції студентів та аспірантів «Сучасні Проблеми Хімії» (Київ, 2019), Київській Конференції з аналітичної хімії «Сучасні Тенденції 2020»

(Київ, 2020), XIII Всеукраїнській науковій конференції студентів та аспірантів «Хімічні Каразінські читання - 2021» (Харків, 2021) та ХХII Міжнародній конференції студентів, аспірантів та молодих вчених «Сучасні Проблеми Хімії» (Київ, 2021).

#### **10. Оцінка мови та стилю дисертації.**

Дисертація є цілісною роботою, написаною науковим стилем мовлення з коректним застосуванням сучасної наукової термінології. Стиль викладення матеріалів дисертації є доступним для сприйняття. Результати досліджень викладені логічно й послідовно. Зміст, структура, оформлення дисертації та кількість публікацій відповідають вимогам «Тимчасового порядку присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 06.03.2019 р. № 167 зі змінами), наказу Міністерства освіти і науки України від 12.01.2017р. № 40 «Про затвердження вимог до оформлення дисертацій».

#### **11. Відповідність змісту дисертації спеціальності з відповідної галузі знань, з якої вона подається до захисту.**

За своїм фаховим спрямуванням, науковою новизною і практичною значимістю дисертаційна робота Бердника М.І. відповідає спеціальності 102 - Хімія. Здобувачем повністю виконано освітню та наукову складову освітньо-наукового рівня вищої освіти.

#### **12. Рекомендація дисертації до захисту.**

Дисертаційна робота Бердника Михайла Ігоровича «Метод L<sub>1</sub>-регуляризації для опису фізико-хімічних властивостей молекул» відповідає вимогам, передбаченим пунктом 10 «Тимчасового порядку присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 06.03.2019 р. № 167 зі змінами).

Враховуючи високий рівень виконаних досліджень, актуальність теми роботи, наукову новизну результатів та їх наукове і практичне значення, рішення фахового семінару кафедри хімічного матеріалознавства хімічного факультету Харківського національного університету імені В. Н. Каразіна, проведеного 30.09.2021 р., дисертація Бердника Михайла Ігоровича «Метод L<sub>1</sub>-регуляризації для опису фізико-хімічних властивостей молекул» рекомендується до захисту в спеціалізованій вченій раді для здобуття ступеня доктора філософії з галузі знань 10 - Природничі науки за спеціальністю 102 - Хімія.

Рецензент

д. ф.-м. н., проф.  
кафедри прикладної хімії,  
Харківського національного  
університету імені В. Н. Каразіна

Владислав ЧЕРАНОВСЬКИЙ

Рецензент

Проректор з науково-педагогічної роботи  
Харківського національного університету  
імені В. Н. Каразіна, к. х. н., доцент,  
доцент кафедри хімічного матеріалознавства

Антон ПАНТЕЛЕЙМОНОВ

07.10.2021

