

Д. М. Клекоць¹, О. А. Безшийко¹, Л. О. Голінка-Безшийко^{1*},
В. Кубицький², І. Чайковська²

¹ Київський національний університет імені Тараса Шевченка, Київ, Україна

² Університет Париж-Сакле, Національний інститут ядерної фізики та фізики елементарних частинок, Лабораторія фізики двох нескінченностей імені Ірен Жоліо-Кюрі, Орсе, Франція

*Відповідальний автор: lyalkagb@gmail.com

ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЗНАЧЕННЯ ПАРАМЕТРІВ ЛІНІЙНОЇ ОПТИКИ ПУЧКА В НАКОПИЧУВАЛЬНОМУ КІЛЬЦІ ThomX

Параметри лінійної оптики є одними серед найважливіших характеристик пучка, що контролюються на прискорювачах заряджених частинок. Класичні методи аналізу, такі як компонентно незалежний аналіз, використовують покази моніторів позиції з кожного оберту пучка. Як альтернатива компонентно незалежному аналізу запропоновано використання машинного навчання та нейронних мереж для визначення параметрів лінійної оптики пучка, при цьому використовуються такі ж вхідні дані, як і для класичного алгоритму. У роботі представлено тренування та використання нейронної мережі для аналізу даних з накопичувального кільця прискорювального комплексу ThomX.

Ключові слова: лінійна оптика пучка, монітори позиції пучка, нейронні мережі, машинне навчання.

1. Вступ

Параметри лінійної оптики пучка є важливими для опису та контролю руху заряджених частинок усередині прискорювача. Також визначення параметрів лінійної оптики пучка є дуже важливим, зокрема для контролю та налаштування прискорювача під час уведення його в експлуатацію. Дані параметри є властивостями прискорювача, отже вони повністю визначаються розташуванням магнітів у прискорювачі та їхнім полем. Вимірювання параметрів лінійної оптики відкриває можливість віднайти невідповідність реальних параметрів магнітів прискорювача по відношенню до референтних значень, для яких прискорювач було спроектовано, тим самим забезпечуючи необхідну інформацію для налаштування прискорювача. Особливо цікавими в цьому плані є методи аналізу, що не використовують попередню інформацію про розташування компонентів прискорювача, що дає змогу суттєвим чином уникнути упередженості в аналізі. Одними з найвизначніших серед таких методів аналізу є PCA [1] та ICA [2, 3], що використовують лише покази моніторів позиції пучка на кожному оберті пучка. Хоча за допомогою даних методів можна ефективно відрізнити різні моди коливань, однак для встановлення їхньої природи потрібно виконувати подальший аналіз. Як альтернатива класичним методам запропоновано використання моделі машинного навчання для визначення параметрів лінійної оптики, зокрема для визначення бета-функції та дисперсії, що характеризують, у лінійному наближенні, область, у якій фізично перебуває пучок.

2. Теорія

Як відомо, в процесі руху пучок здійснює бетатронні та синхротронні коливання навколо замкненої траєкторії. Вимірювання положення пучка на кожному оберті здійснюється за допомогою моніторів позиції пучка. Розглянемо коливання пучка у горизонтальній площині, позначимо відхилення від замкненої орбіти як x . Тоді, враховуючи бетатронні та синхротронні коливання, можемо представити відхилення від усередненої орбіти, в i -му моніторі та на j -му оберті у вигляді

$$x_{ij} = \sqrt{\epsilon\beta_i} \cos(2\pi\nu_b j + \mu_{b,i}) + D_i \delta \cos(2\pi\nu_s j + \mu_{s,i}). \quad (1)$$

Тут ν_b та ν_s є відповідно частотами бетатронних та синхротронних коливань. Параметрами пучка є поперечний еммітанс (у горизонтальній площині) ϵ та відносне відхилення імпульсу від референтного значення $\delta = \frac{p - p_0}{p_0}$. Зсув бетатронної та синхротронної фази, у положенні i -го монітора, відповідно позначені $\mu_{b,i}$ та $\mu_{s,i}$. Лінійними параметрами пучка у i -му моніторі є бета-функція β_i та дисперсія D_i . Як видно з рівняння, амплітудами коливань є безпосередньо $A_{b,i} = \sqrt{\epsilon\beta_i}$ та $A_{s,i} = D_i \delta$, для бетатронних та синхротронних коливань відповідно. Числові значення параметрів лінійної оптики можуть бути визначені з відповідних формул для амплітуди.

© Д. М. Клекоць, О. А. Безшийко, Л. О. Голінка-Безшийко, В. Кубицький, І. Чайковська, 2024

Як вже згадувалося у вступі, одним серед відомих компонент незалежних методів визначення параметрів пучка є PCA [1] та, як покращена версія, ICA [2, 3]. Основна ідея цих методів полягає у розкладенні коливань позиції пучка, вимірних з послідовних показів моніторів, на відповідні амплітудні та сигнальні множники. Якщо подати покази моніторів позиції пучка у вигляді матриці $\hat{X} = \|x_{ij}\|$ з кількістю рядків, що дорівнює кількості моніторів, та кількістю стовбців, що дорівнює кількості послідовних вимірювань, то алгоритми PCA та ICA можна коротко описати як розклад матриці \hat{X} у вигляді

$$\hat{X} = \hat{A} \cdot \hat{s}, \quad (2)$$

де матриця \hat{A} відповідає за амплітудні характеристики коливань та їхню просторову структуру. Матриця \hat{s} представляє часові характеристики коливань. Такий розклад забезпечує однакові часові характеристики коливань (тобто затухання та частота будуть однакові при виборі певної моди коливань, наприклад бетатронної), однак амплітуди та фази є різними для кожного монітора.

Хоча метод ICA є доволі потужним у розділенні мод коливань, не завжди вдається виконувати аналіз автоматично, оскільки порядок мод у розділених коливаннях є довільним, тобто потрібно додатково, інколи в ручному режимі, встановлювати яке коливання належить до синхротронних, а яке до бетатронних. Даного недоліку позбавлений запропонований авторами альтернативний метод використання нейронної мережі для визначення амплітуд бетатронних та синхротронних коливань. Основна перевага запропонованого методу є в тому, що в процесі тренування моделі використовуються референтні дані, в яких бетатронні та синхротронні коливання розділені, таким чином натренована модель навчається їх розділяти при передбаченні.

Нейронна мережа, використувана авторами, була розроблена для передбачення амплітуд коливань у накопичувальному кільці прискорювального комплексу ThomX [4], який під час розробки моделі перебував на стадії введення в експлуатацію [5]. Накопичувальне кільце ThomX складається з чотирьох симетричних частин та містить 12 моніторів позиції пучка. Слід зазначити, що запропонована методика може бути використана і для інших прискорювачів.

Таблиця 1. Послідовність шарів, що використовувалися нейронною мережею, та їхні характеристики

№		Шар нейронної мережі	Розмірність вихідного тензора
1	Згорткова частина нейронної мережі	Вхідний шар	12 × 201 × 1
2		Згортковий шар з ядром розмірності 1 × 5, функцією активації gelu та 32 фільтрами	12 × 197 × 32
3		Шар усередненого об'єднання з ядром 1 × 3	12 × 65 × 32
4		Згортковий шар з ядром розмірності 1 × 3, функцією активації gelu та 64 фільтрами	12 × 63 × 64
5		Шар усередненого об'єднання з ядром 1 × 3	12 × 21 × 64
6		Згортковий шар з ядром розмірності 1 × 3, функцією активації gelu та 128 фільтрами	12 × 19 × 128
7		Шар усередненого об'єднання з ядром 1 × 3	12 × 6 × 128
8		Шар зміни розмірності тензора	9216
9	Повнозв'язна частина	Повнозв'язний шар з 960 нейронами та функцією активації softplus.	960
10		Повнозв'язний шар з 24 нейронами та функцією активації softplus.	24

В якості вхідних даних алгоритми PCA та ICA використовують матрицю \hat{X} , кожен рядок якої представляє сумарні коливання позиції пучка у місці розташування певного монітора. Однак, для нейронних мереж набагато зручнішим є використання дискретного Фур'є-спектра відповідних значень рядків матриці \hat{X} , оскільки в такому разі коливання з певною частотою представляється

піком на спектрі, що значно покращує розділення різних коливань. Використовувана в даній роботі нейронна мережа приймала в якості вхідних значень модулі від відповідних комплексних значень дискретного Фур'є-спектра. Для реалізації моделі нейронної мережі було використано обчислювальний код TensorFlow [6]. В якості вхідних параметрів модель приймала 201 абсолютне значення дискретного Фур'є-спектра для кожного з

12 моніторів. Варто зазначити, що аналіз проводиться для 400 обертів пучка, однак лише 201 компонент дискретного Фур'є-перетворення є незалежними, саме вони були використані як вхідні дані. На виході модель робить передбачення 24 значень, що є амплітудами синхротронних та бетатронних коливань для 12 моніторів. Використовувана модель складалася з послідовно з'єднаних шарів, характеристики яких представлені у табл. 1. Перша частина моделі представ-

лена згортковими рівнями, де відбувається виділення важливих характеристик сигналу. Другу частину складають повнозв'язні рівні.

3. Тренування нейронної мережі

Тренування здійснювалося на даних, згенерованих псевдовипадковим чином, з рівномірним розподілом відхилень параметрів коливань від модельних значень у межах, зазначених у табл. 2.

Таблиця 2. Межі відхилень параметрів коливань позиції пучка, в яких було згенеровано тренувальні дані

Параметр	Синхротронні коливання	Бетатронні коливання
	Межі відхилення від референтних значень, %	
Частота ν [1/оберт]	± 10	± 10
Зсув фази μ	$\pm 2,5$	± 5
Амплітуда A	± 20	± 20

Положення пучка було згенеровано відповідно до формули (1), що представляло вхідні значення для тренування моделі, референтні значення передбачення моделі (які є амплітудами коливань у тренувальних даних) використовувалися при генеруванні відповідних коливань, отже були відомими для тренування моделі. Коливання, згенеровані псевдовипадковим чином, були віднормовані на максимальне значення відхилення, на таку ж величину були віднормовані і референтні значення. Використане нормування забезпечує масштабованість моделі. Після нормування у згенеровані дані було додано шум, розподілений за Гауссівським розподілом, параметр σ якого лінійно зростає від 0,01 (для першого оберту пучка) до 0,025 (для 400 оберту пучка). Зазначимо, що параметр σ є безрозмірним, оскільки шуми були додані до вже віднормованих даних.

4. Результати

Тренування нейронної мережі здійснювали на наборі з 9600 зразків даних; 2400 зразків вико-

ристовувалися для валідації. Загалом 480 тренувальних епох було пройдено з використанням оптимізатора Adam [7] та функції втрат, що наведена у формулі (3)

$$L = 2 \frac{1}{24} \sum_{i=1}^{24} \left(\frac{\tilde{Y}_i - Y_i}{\tilde{Y}_i} \right)^2. \quad (3)$$

У даній формулі підсумовування йде за всіма 24 вихідними значеннями моделі, Y_i позначають передбачення моделі, \tilde{Y}_i – контрольні значення. Така функція втрат забезпечує квадратичну залежність від різниці референтного та передбаченого значень, а також безпосереднє відхилення нормується на саме референтне значення, що забезпечує рівноцінність у функції втрат показів моніторів, у яких відповідні амплітуди коливань є низькими порівняно з моніторами, в яких амплітуди приймають більше числове значення. Залежність функції втрат для тренувальних та валідаційних даних зображено на рис. 1.

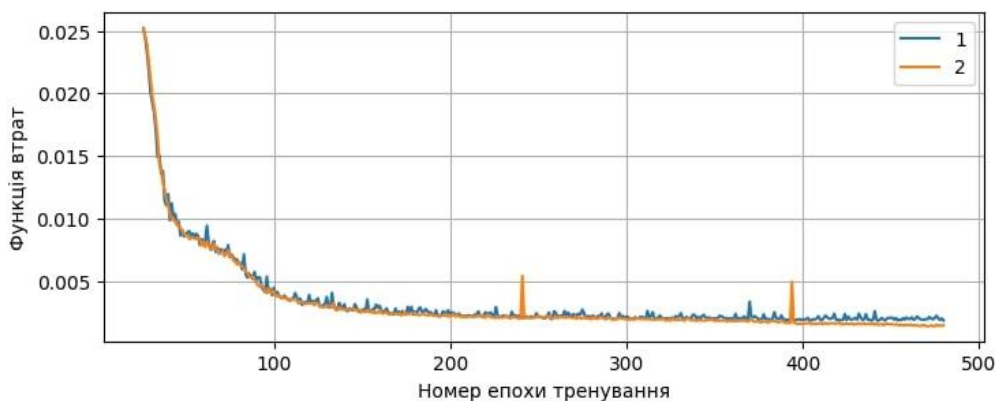


Рис. 1. Залежність від епохи тренування функції втрат, розрахованої на валідаційному (крива 1) та тренувальному (крива 2) наборі даних. (Див. кольоровий рисунок на сайті журналу.)

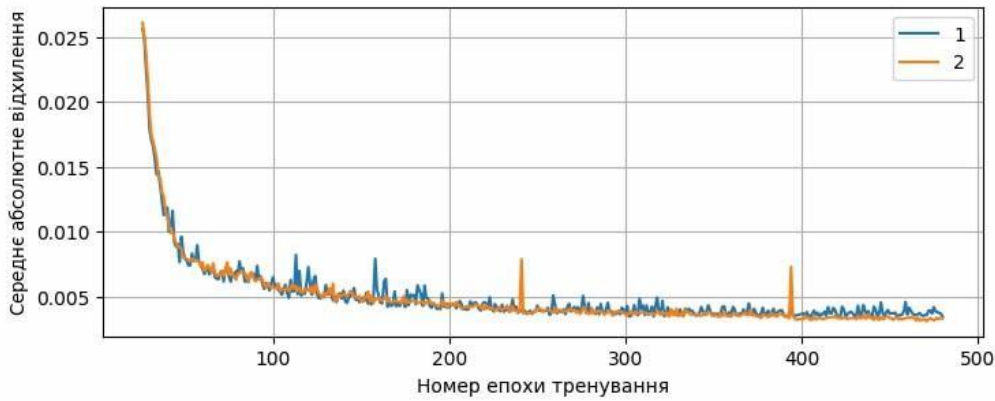


Рис. 2. Залежність від епохи тренування середнього абсолютного відхилення, розрахованого на валідаційному (крива 1) та тренувальному (крива 2) наборі даних. (Див. кольоровий рисунок на сайті журналу.)

В якості метрики було використано середнє абсолютне відхилення значення, передбаченого моделлю, від відповідного контрольного. Залежність середнього абсолютного відхилення від номера епохи тренування зображено на рис. 2. Як видно з рисунка, значення середнього абсолютного відхилення, розраховане на валідаційному наборі даних, виходить на плато після 400 епох, тому подальше тренування не має особливого сенсу. Також з рис. 1 видно, що приблизно після 400 епохи тренування, функція втрат, розрахована на тренувальному наборі даних, продовжує дещо зменшуватися, в той час, як функція втрат, розрахована на валідаційному наборі, не зменшу-

ється, що є індикатором того, що модель починає затреновуватися під специфічний тренувальний набір даних.

У результаті тренування була відібрана модель з найменшим середнім абсолютним відхиленням, що було розраховане на валідаційному наборі даних та становило $3,4 \cdot 10^{-3}$. Метрики відібраної моделі було перевірено на тестовому наборі даних та становили такі значення:

- функція втрат, що розраховувалася за формулою (3) становила $9,35 \cdot 10^{-4}$;
- середнє абсолютне відхилення становило $3,42 \cdot 10^{-3}$.

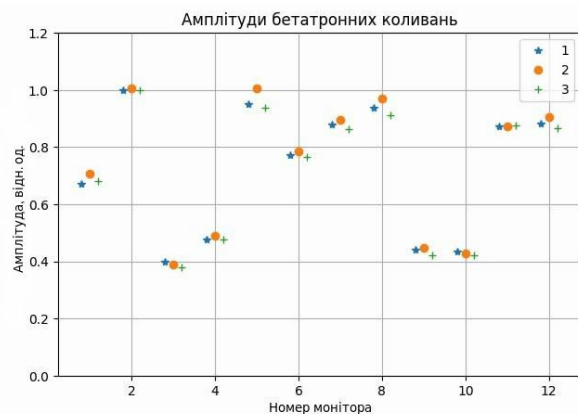
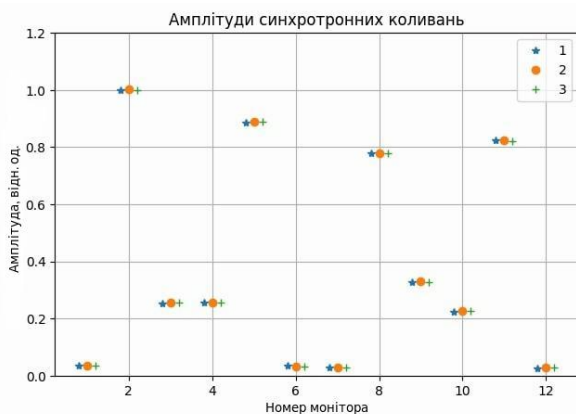


Рис. 3. Амплітуди синхротронних та бетатронних коливань у кожному моніторі. Наведено референтні величини (точки 1), величини реконструйовані за допомогою нейронної мережі (точки 2) та реконструйовані за допомогою ІСА аналізу (точки 3). Значення на кожному графіку було віднормовано на максимальне значення амплітуди. (Див. кольоровий рисунок на сайті журналу.)

На рис. 3 зображено, в якості демонстрації, контрольні значення амплітуд тестувального набору даних, а також передбачення моделі та передбачення з використанням методу ІСА для одного зразка з тестувального набору даних. Загалом визначення амплітуд синхротронних коливань є точнішим, це пов'язано з тим, що бетатронні коливання мають загалом менші амплітуди порівняно з синхротронними коливаннями,

тобто вони більш чутливі до шумів, а також бетатронні коливання менш помітні на фоні синхротронних.

5. Висновки

У даній роботі запропоновано використання нейронної мережі та методів машинного навчання для визначення амплітуд бетатронних та синхро-

тронних коливань, через які можуть бути визначені параметри лінійної оптики пучка, зокрема дисперсія та бета-функція в моніторах позиції пучка. В якості вхідних даних для розрахунку використовуються послідовні вимірювання позиції пучка на кожному оберті та у кожному з моніторів позиції пучка.

Методика продемонстрована на прикладі накопичувального кільця прискорювального комплексу ThomX. Для тренування нейронної мережі використовувалися дані, згенеровані псевдовипадковим чином з відхиленнями від номінальних значень параметрів коливань, для яких було спроектовано накопичувальне кільце. Тренування здійснювалося протягом 480 епох з 9600 зразків тренувальних даних та 2400 валідаційних даних. Модель з найменшим середнім абсолютним відхиленням передбачених моделлю значень від контрольних значень була відібрана як найкраща

в процесі тренування, та перевірена на тестувальному наборі даних. Найкраще натренована модель була перевірена на тестувальному наборі даних, що складав 12000 зразків даних, середнє абсолютне відхилення передбачених значень від контрольних становило $3,42 \cdot 10^{-3}$, при нормуванні контрольних значень на максимальну амплітуду.

Наукова робота авторів з Київського національного університету імені Тараса Шевченка і Лабораторії лінійних прискорювачів (Франція) проводилася в рамках міжнародної асоційованої лабораторії LIA IDEATE.

Дослідження частково підтримане фінансуванням НФДУ в рамках проекту «Підвищення якості та безпеки при проведенні променевої терапії онкологічних захворювань і рентгенодіагностики» за реєстраційним номером 2021.01/0211 (конкурс «Наука для безпеки та сталого розвитку України»).

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ / REFERENCES

1. X. Huang. *Principal component analysis (PCA)*. In: *Beam-Based Correction and Optimization for Accelerators* (Taylor & Francis Group, 2020) p. 131.
2. X. Huang et al. Application of independent component analysis to Fermilab Booster. *Phys. Rev. ST Accel. Beams* 8 (2005) 064001.
3. X. Yang, X. Huang. A method for simultaneous linear optics and coupling correction for storage rings with turn-by-turn beam position monitor data. *Nucl. Instrum. Methods A* 828 (2016) 97.
4. A. Variola et al. *ThomX Technical Design Report*. LAL / RT 14-21. *Laboratoire de l'accélérateur linéaire* (2014). 164 p.
5. V. Kubytskyi et al. Commissioning of the ThomX Storage Ring. *J. Phys.: Conf. Ser.* 2687 (2024) 032031.
6. M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems. *arXiv:1603.04467v2 [cs.DC]* (2016).
7. D.P. Kingma, J. Ba. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980v9 [cs.LG]* (2017).

D. Klekots¹, O. Bezshyyko¹, L. Golinka-Bezshyyko^{1,*}, V. Kubytskyi², I. Chaikovska²

¹ *Taras Shevchenko National University of Kyiv, Kyiv, Ukraine*

² *University Paris-Saclay, National Institute for Nuclear and Particle Physics, Laboratory of the Physics of the Two Infinities Irène Joliot-Curie, Orsay, France*

*Corresponding author: lyalkagb@gmail.com

APPLYING THE MACHINE LEARNING METHODS TO DETERMINE THE LINEAR OPTICS PARAMETERS IN THE ThomX COLLECTOR RING

The linear optics parameters are one of the most significant properties of the beam, which are controlled at the particle accelerators. Classical methods of analysis, such as component-independent analysis, employ turn-by-turn readings of the beam position monitors. As an alternative to the component-independent analysis, machine learning and neural networks are proposed for determining the beam parameters. This approach relies on the same input data as classical algorithms. This work shows training and usage of the neural network for analysis of the data from the collector ring of the ThomX accelerator facility.

Keywords: linear beam optics, beam position monitor, neural networks, machine learning.

Надійшла / Received 17.10.2024