

Научная статья

УДК 539.12

DOI: <https://doi.org/10.18721/JPM.16415>

МОДЕЛИРОВАНИЕ ПОЛУИНКЛЮЗИВНОГО, ГЛУБОКОНЕУПРУГОГО РАССЕЙЯНИЯ ЛЕПТОНА НА ПРОТОНЕ ПРИ ЭНЕРГИЯХ 20 – 100 ГэВ НА ОСНОВЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНОЙ НЕЙРОННОЙ СЕТИ

А. А. Лобанов ✉, **Я. А. Бердников**

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Россия

✉ lobanov2.aa@edu.spbstu.ru

Аннотация. Данная работа продолжает цикл статей, посвященных развитию возможностей генератора событий глубоконеупругого лептон-протонного рассеяния на основе генеративно-сопоставительной сети (ГСС). Здесь рассмотрены полуинклюзивные реакции глубоконеупругого рассеяния с регистрацией адрона. Показано, что ГСС позволяет с высокой точностью генерировать распределения физических характеристик конечных лептона и адрона в диапазоне начальных энергий 20 – 100 ГэВ в системе центра масс.

Ключевые слова: полуинклюзивное, глубоконеупругое рассеяние; машинное обучение; нейронная сеть; генеративно-сопоставительная сеть

Для цитирования: Лобанов А. А., Бердников Я. А. Моделирование полуинклюзивного, глубоко неупругого рассеяния лептона на протоне при энергиях 20 – 100 ГэВ на основе генеративно-сопоставительной нейронной сети // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 4. С. 189–197. DOI: <https://doi.org/10.18721/JPM.16415>

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

© Лобанов А. А., Бердников Я. А., 2023. Издатель: Санкт-Петербургский политехнический университет Петра Великого.

Original article

DOI: <https://doi.org/10.18721/JPM.16415>

SIMULATION OF SEMI-INCLUSIVE DEEP INELASTIC LEPTON SCATTERING ON A PROTON AT ENERGIES OF 20 – 100 GeV ON THE BASIS OF A GENERATIVE-ADVERSARIAL NEURAL NETWORK

A. A. Lobanov ✉, **Ya. A. Berdnikov**

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ lobanov2.aa@edu.spbstu.ru

Abstract. This paper continues a series of articles devoted to developing the capabilities of a deep inelastic lepton-proton scattering event generator based on the generative adversarial network (GAN). The investigation has focused on semi-inclusive reactions of deep inelastic scattering and, particularly, on hadron registration. The results confirmed that GAN could accurately generate distributions of physical properties of leptons and hadrons. It worked for

different types of leptons and hadrons in the range of initial energies from 20 to 100 GeV in the center-of-mass system. The GAN demonstrated the possibility to preserve the inherent correlation between the characteristics of leptons and protons.

Keywords: semi-inclusive deep inelastic scattering, machine learning, neural network, generative-adversarial network

For citation: Lobanov A. A., Berdnikov Ya. A., Simulation of semi-inclusive deep inelastic lepton scattering on a proton at energies of 20 – 100 GeV on the basis of a generative-adversarial neural network, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (4) (2023) 189–197. DOI: <https://doi.org/10.18721/JPM.16415>

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Как известно [1], в современных экспериментальных исследованиях в области физики высоких энергий приходится иметь дело со все большими и большими массивами данных. Источниками последних выступают крупномасштабные эксперименты или результаты моделирования. Работа с этой информацией требует привлечения больших вычислительных мощностей и временных затрат.

В качестве одного из подходов для решения вышеупомянутых проблем могут быть использованы методы машинного обучения [2]. На основе этих методов можно построить программы компьютерного моделирования (называемые генераторами событий), открывающие следующие новые возможности:

на основе экспериментальных результатов по изучению продуктов взаимодействия частиц и ядер в дискретных точках предсказывать по значениям начальной энергии характеристики вторичных частиц при любых энергиях в исследуемом интервале на основе интерполяции (и, возможно, экстраполяции), причем быстро и без больших вычислительных затрат;

даже при отсутствии результатов экспериментов, для разработки вышеупомянутых программных продуктов можно использовать результаты моделирования исследуемых взаимодействий частиц и ядер, полученные на основе метода Монте-Карло [3].

В статье [4] было рассмотрено применение генеративно-сопоставительной сети (ГСС) для создания генератора инклюзивного, глубоко неупругого лептон-протонного рассеяния.

Данная работа продолжает развивать эту проблему; теперь рассмотрено распространение возможностей указанного генератора событий [4] на полуинклюзивное, глубоко неупругое рассеяние с регистрацией адрона.

Целью работы является построение генератора, который может обучаться на экспериментальных (или полученных в результате компьютерного моделирования) данных и позволяет на основе интерполяции и экстраполяции получать промежуточные данные, поскольку эксперимент невозможно проводить при любых значениях начальной энергии.

Интерес к полуинклюзивным процессам вызван несколькими причинами.

Во-первых, регистрация дополнительного адрона позволяет больше узнать о структуре протона. Так, тип рожденного в лептон-протонном взаимодействии адрона будет зависеть от того, с каким ароматом кварка в протоне провзаимодействовал виртуальный фотон, испущенный заряженным лептоном [5].

Во-вторых, характеристики дополнительного адрона могут нести информацию о процессах адронизации партонов [5].

В-третьих, в ходе полуинклюзивных процессов есть возможность измерять различные спиновые и азимутальные асимметрии, а это позволяет составить представление о спиновой структуре протона [6].

Методика исследования

Характеристиками конечного состояния заряженного лептона (e^+ , e^- , μ^+ , μ^-) и адрона (π^0 , π^+ , π^- , K^+ , K^-) выступают их 4-импульсы $p_l = (E_l, \mathbf{p}_l)$ и $p_h = (E_h, \mathbf{p}_h)$ соответственно, где E_l – полная энергия рассеянного лептона; p_l , \mathbf{p}_l – четырех- и трехмерный векторы



импульса лептона, причем последний определяется через его компоненты p_{xl}, p_{yl}, p_{zl} ; E_h – полная энергия адрона, p_h, \mathbf{p}_h – четырех- и трехмерный векторы импульса адрона, и также компоненты последнего – p_{xh}, p_{yh}, p_{zh} .

Для того, чтобы ГСС могла предсказывать 4-импульс различных адронов ($\pi^0, \pi^+, \pi^-, K^+, K^-$), их тип (как и тип лептона) передается на вход ГСС в качестве дополнительных параметров вместе с начальной энергией E_0 , определяемой как $E_0 \approx \sqrt{s_{IN}}/2$, где $\sqrt{s_{IN}}$ – начальная энергия в системе центра масс лептон-протон [4].

Поскольку в настоящее время получать характеристики конечных лептонов и адронов из эксперимента не представляется возможным (из-за отсутствия экспериментов), конечные состояния лептонов и адронов были получены с использованием программного пакета RUTHIA8 [7].

Для каждого типа лептона (e^+, e^-, μ^+, μ^-) и адрона ($\pi^0, \pi^+, \pi^-, K^+, K^-$) было сгенерировано по 100 тыс. событий при начальных энергиях $\sqrt{s_{IN}} = 20, 40, 60, 80$ и 100 ГэВ. Из каждого события были получены значения 4-импульсов конечного лептона и адрона (эталонные данные).

Для решения проблем, связанных с нерегулярностями в распределениях по величинам E_l, E_h и p_{zl} , в настоящем исследовании, как и в работе [4], использована генерация не самих величин E_l, E_h, p_{zl} , а величин, полученных в результате их преобразования (преобразованные величины):

$$T(p_{zl}) = \log[(E_0 - p_{zl})/(1 \text{ GeV}/c)],$$

$$T(E_l) = \log[(E_0 - E_l)/(1 \text{ GeV}/c)],$$

$$T(E_h) = \log[E_h/(1 \text{ GeV}/c)].$$

Распределение по преобразованным величинам, как показано в работе [4], становится более гладким и позволяет избежать предсказания нефизических значений.

Так же, как в статье [4], в данной работе генератор событий строится на основе ГСС с функцией потерь по методу наименьших квадратов [8].

Генератор состоит из 5 слоев по 512 нейронов с функцией активации «Leaky ReLU» и показателем 0,2 [9]. Ему на вход поступают 128-мерный вектор шума (вектор значений, полученных из распределения Гаусса со средним, равным 0, и дисперсией, равной 1), энергия E_0 , тип лептона и тип адрона. На выходе генератора получаем 8 характеристик:

$$p_{xl}, p_{yl}, T(p_{zl}), T(E_l), p_{xh}, p_{yh}, p_{zh} \text{ и } T(E_h),$$

соответствующих лептону и адрону.

На основе этих характеристик в модели рассчитываются дополнительные величины, используемые для повышения точности предсказания ГСС, [4]:

$p_{Tl} = \sqrt{p_{xl}^2 + p_{yl}^2}$, $p_{Th} = \sqrt{p_{xh}^2 + p_{yh}^2}$ – поперечные импульсы лептона и адрона, соответственно;

$\varphi_l = \arctan(p_{zl}/p_{Tl})$, $\varphi_h = \arctan(p_{zh}/p_{Th})$ – азимутальные углы лептона и адрона, соответственно;

$\theta_l = \arctan(p_{yl}/p_{xl})$, $\theta_h = \arctan(p_{yh}/p_{xh})$ – полярные углы лептона и адрона, соответственно.

Все дополнительные величины далее передаются на вход дискриминатора при обучении.

Дискриминатор также состоит из 5 слоев по 512 нейронов с функцией активации «Leaky ReLU» и показателем 0,2 [9]. Для борьбы с переобучением дискриминатора [10], к каждому слою применяется «dropout layer» с коэффициентом 10 % [11], который случайным образом обнуляет 10 % весов слоя. Для более стабильного обучения, ко всем слоям дополнительно применяется спектральная нормализация [12]. Выходной слой состоит из одного нейрона с линейной функцией активации. Чем больше полученное значение, тем «увереннее» дискриминатор считает рассмотренные значения «реалистичными».

Модель обучалась 400 эпох. Оптимизатор градиентного спуска был выбран RMSProp со значениями $\rho = 0,9$ [13], шагов обучения $1 \cdot 10^{-4}$ для генератора и $5 \cdot 10^{-5}$ для

дискриминатора. Использование разных шагов обучения способствует лучшей сходимости при обучении, что показано в работе [14].

В качестве критерия оценки степени расхождения между эталонными данными и сгенерированными ГСС было использовано расстояние Кульбака – Лейбера (КЛ) [15]. Данный критерий применялся для сравнения гистограмм полученных распределений. В этом случае расстояние Кульбака – Лейбера DKL определяется следующим образом [15]:

$$D_{KL}(P \parallel Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i},$$

где P, Q – распределения эталонных и сгенерированных данных, соответственно; p_i, q_i – вероятности i -х бинов гистограмм эталонных и сгенерированных данных; n – число бинов.

Результаты моделирования

По причине большого числа различных вариантов рассеяния (это разные типы лептонов и адронов, а также разные значения начальной энергии E_0), для демонстрации результатов предсказания ГСС далее приведены лишь отдельные возможные варианты.

На рис. 1 представлены распределения по величинам p_T, θ, φ для позитрона e^+ и отрицательного каона K^- , полученные с помощью ГСС и РУТНИА8. Под множественностью (Multiplicity) понимается (на рис. 1 и далее) число отсчетов в бине, нормированное на общее число событий. Видно, что модель генерирует величины, распределения которых практически идентичны, о чем говорят приведенные на графиках значения расстояния Кульбака – Лейбера, а также логарифмы отношения предсказаний ГСС к данным РУТНИА8, приведенные для каждого графика.

На рис. 2 представлены распределения величин p_T, θ, φ для мюона μ^- и положительного каона K^+ , полученные с помощью ГСС и РУТНИА8. Приведенные данные

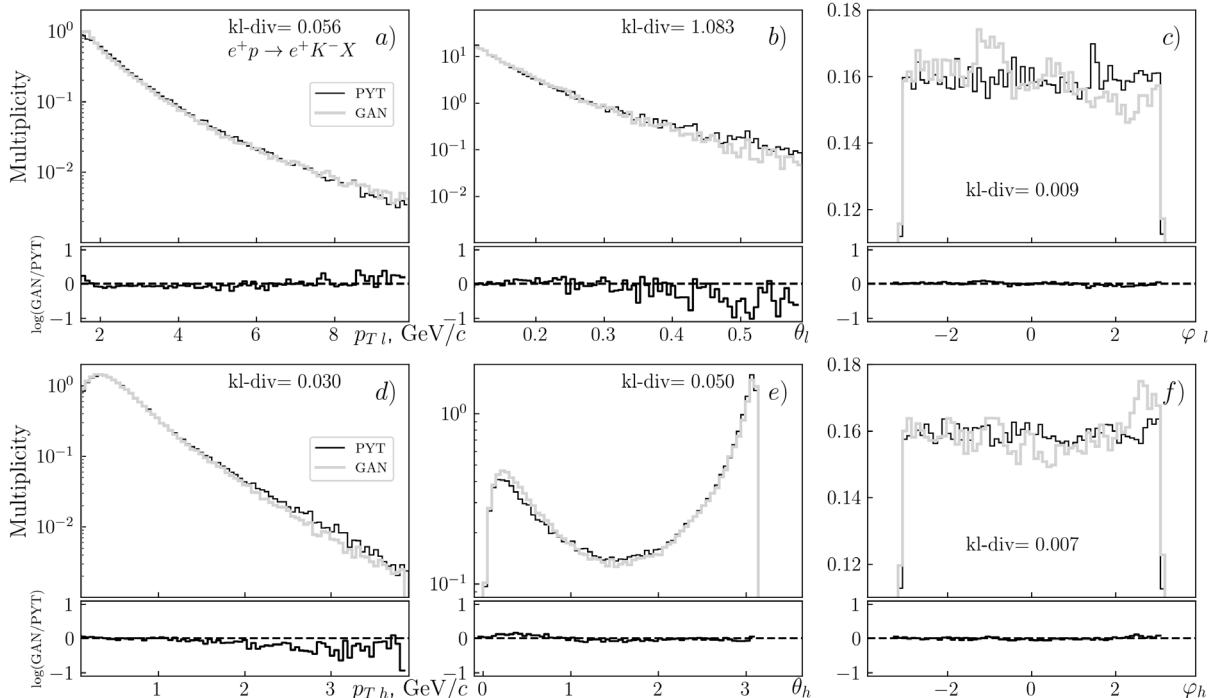


Рис. 1. Распределения по величинам p_T, θ, φ для позитронов e^+ (a, b, c) и отрицательных каонов K^- (d, e, f) при начальной энергии $E_0 = 50$ ГэВ.

Данные получены с помощью ГСС (кривые серого цвета) и РУТНИА8 (черного цвета). Для каждого распределения приведено соответствующее значение расстояния КЛ (kl-div) и график логарифма отношения предсказания ГСС к РУТНИА8 (GAN/PYT)

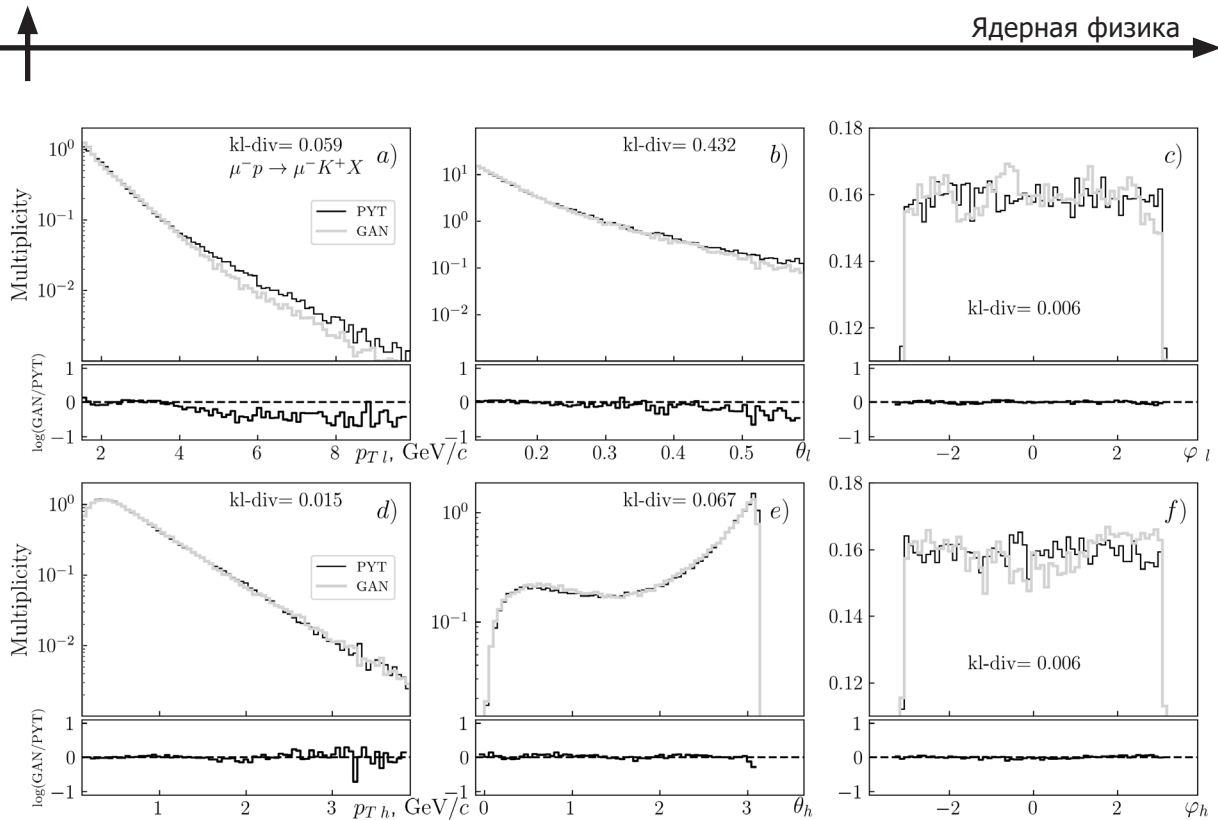


Рис. 2. Графики, аналогичные приведенным на рис. 1, но для мюонов μ^- (a, b, c) и положительных каонов K^+ (d, e, f) при начальной энергии $E_0 = 20$ ГэВ

демонстрируют, что модель способна так же точно работать с различными лептонами и адронами при разных начальных энергиях.

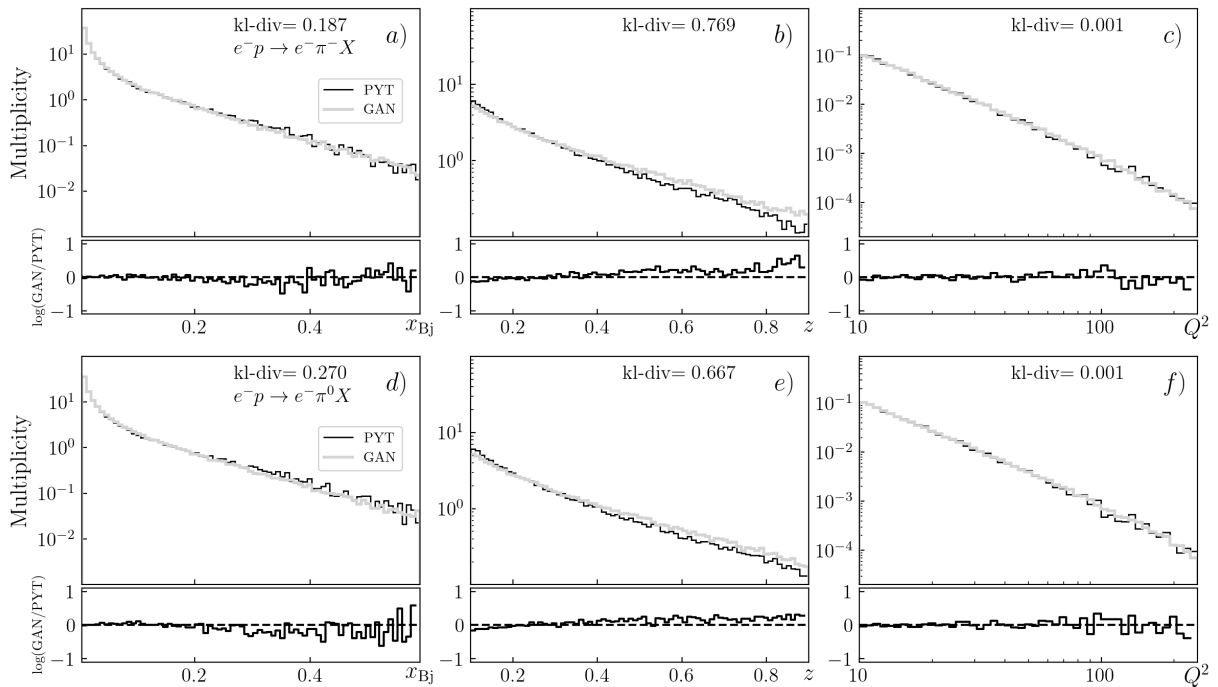


Рис. 3. Распределения по величинам x_{Bj} , z , Q^2 для реакций $e^-p \rightarrow e^- \pi^- X$ (a, b, c) и $e^-p \rightarrow e^- \pi^0 X$ (d, e, f) соответственно, при начальной энергии $E_0 = 40$ ГэВ.

Для каждого распределения приведено соответствующее значение расстояния КЛ (kl-div) и график логарифма отношения предсказания ГСС к РYTHIA8 (GAN/PYT)

На рис. 3 представлены распределения величин квадрата переданного импульса $Q^2 = -q^2$ (q – импульс виртуального фотона), а также переменной Бьёркина $x_{\text{Bj}} = Q^2/2Pq$ (P – импульс налетающего протона) и доли энергии виртуального фотона, переданной адрону, $z = P \cdot P_h / P \cdot q$ (P_h – импульс адрона) для ядерных реакций

$$e^-p \rightarrow e^- \pi X \text{ и } e^-p \rightarrow e^- \pi^0 X,$$

где буквой X обозначены все остальные продукты реакции.

Из представленных результатов следует, что распределения, сгенерированные моделью, различаются слабо; на это указывают значения расстояния КЛ, полученные для каждого распределения.

На рис. 4 приведены распределения по величинам x_{Bj} , z , Q^2 для реакций $e^+p \rightarrow e^+ \pi^+ X$ и $e^+p \rightarrow e^+ K^+ X$. Анализ полученных данных приводит к заключению, что точность предсказания ГСС относительно эталонных данных РYTHIA8 сохраняется при различных типах лептона и адрона и разных значениях начальной энергии.

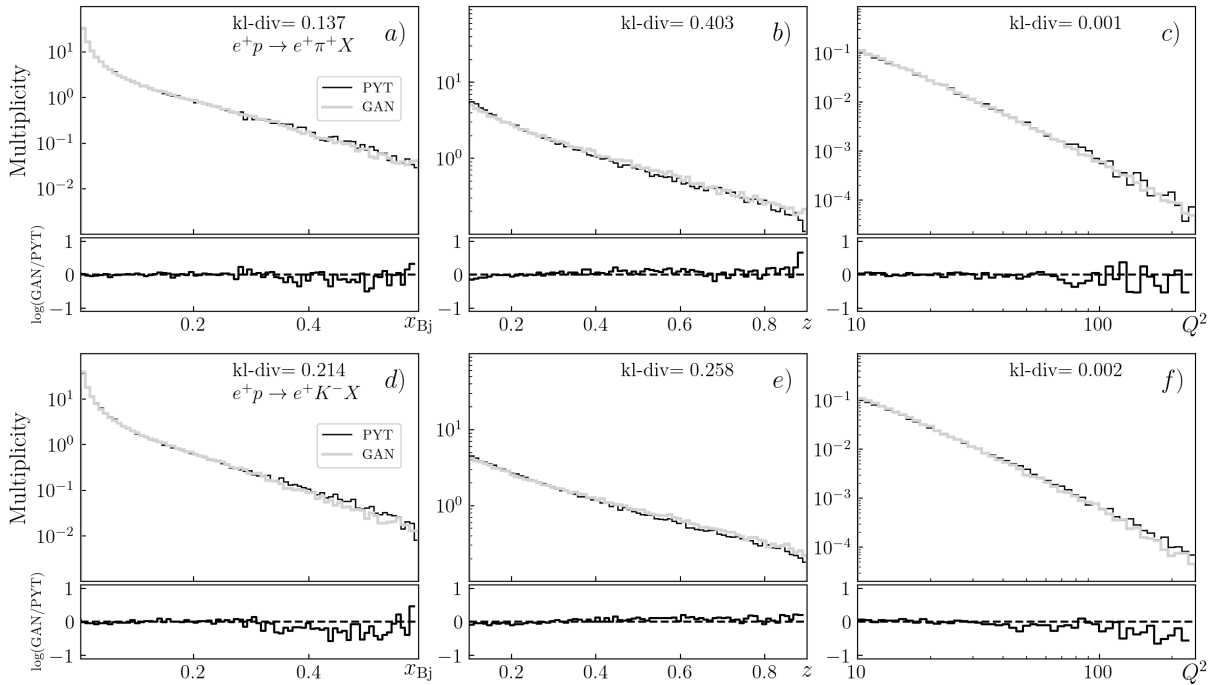


Рис. 4. Распределения по величинам x_{Bj} , z , Q^2 для реакций $e^-p \rightarrow e^- \pi X$ (a, b, c) и $e^-p \rightarrow e^- \pi^0 X$ (d, e, f) при начальной энергии $E_0 = 30$ ГэВ.

Для каждого распределения приведено соответствующее значение расстояния КЛ (kl-div) и график логарифма отношения предсказания ГСС к РYTHIA8 (GAN/PYT)

Заключение

В данной статье разработана модель генеративно-состязательной сети, способная предсказывать характеристики конечных лептонов (e^+ , e^- , μ^+ , μ^-) и адронов (π^0 , π^+ , π^- , K^+ , K^-) в полуинклюзивном лептон-протонном, глубоконеупругом рассеянии в диапазоне начальных энергий 20 – 100 ГэВ.

Установлено, что вышеупомянутая модель ГСС способна с высокой точностью получать компоненты 4-импульсов конечных лептонов и адронов.

Показано, что полученная модель способна с высокой точностью рассчитывать распределения для частиц: по поперечному импульсу частиц p_T , по азимутальному (ϕ) и полярному (θ) углам, по переменной Бьёркена x_{Bj} , по значениям доли энергии виртуального фотона z и квадрату переданного лептоном адрону Q^2 . Распределения по этим величинам показывают высокую точность относительно эталонных данных, что говорит о способности модели сохранять внутренние связи между величинами.



Установлено также, что модель ГСС точно предсказывает характеристики лептонов и адронов как при начальных энергиях, при которых велось обучение, так и при интерполированных значениях начальной энергии (промежуточные значения энергии).

СПИСОК ЛИТЕРАТУРЫ

1. **Cremonesi M., Bellini C., Bian B., et al.** Using big data technologies for HEP analysis. arXiv: 1901.07143, 2019. <https://doi.org/10.48550/arXiv.1901.07143>
2. **Jordan M. I., Mitchell T. M.** Machine learning: Trends, perspectives, and prospects // *Science*. 2015. Vol. 349. No. 6245. Pp. 255–260.
3. **Weinzierl S.** Introduction to Monte Carlo methods. arXiv: 0006269, 2000. <https://doi.org/10.48550/arXiv.hep-ph/0006269>.
4. **Лобанов А. А., Бердников Я. А.** Генератор глубокого неупругого рассеяния лептонов на протоне на основе генеративно-состязательной нейронной сети // *Научно-технические ведомости СПбГПУ. Физико-математические науки*. 2023. Т. 16. № 4. С. 181–188.
5. **Aschenauer E. C., Borsa I., Sassot R., Van Hulse C.** Semi-inclusive deep-inelastic scattering, parton distributions, and fragmentation functions at a future electron-ion collider // *Physical Review D*. 2019. Vol. 99. No. 9. P. 094004.
6. **Barone V., Boglione M., Hernandez J. O. G., Melis S.** Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering // *Physical Review D*. 2015. Vol. 91. No. 7. P. 074019.
7. **Sjöstrand T., Mrenna S., Skands P.** A brief introduction to PYTHIA 8.1 // *Computer Physics Communications*. 2008. Vol. 178. No. 11. Pp. 852–867.
8. **Mao X., Li Q., Xie H., Lau R. Y. K., Wang Zh., Smolley S. P.** On the effectiveness of least squares generative adversarial networks // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2019. Vol. 41. No. 12. Pp. 2947–2960.
9. **Sharma O.** A new activation function for deep neural network // *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. IEEE, Faridabad, India, February 14–16, 2019. Pp. 84–86.
10. **Hawkins D. M.** The problem of overfitting // *Journal of Chemical Information and Computer Sciences*. 2003. Vol. 44. No. 1. Pp. 1–12.
11. **Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R.** Dropout: A simple way to prevent neural networks from overfitting // *The Journal of Machine Learning Research*. 2014. Vol. 15. Pp. 1929–1958.
12. **Miyato T., Kataoka T., Koyama M., Yoshida Y.** Spectral normalization for Generative Adversarial Networks. arXiv: 1802.05957/v1, 2018. <https://doi.org/10.48550/arXiv.1802.05957>.
13. **Xu D., Zhang Sh., Zhang H., Mandic D. P.** Convergence of the RMSProp deep learning method with penalty for nonconvex optimization // *Neural Networks*. 2021. Vol. 139. July. Pp. 17–23.
14. **Heusel M., Ramsauer H., Unterthiner T., Nessler B., Hochreiter S.** GANs trained by a two time-scale update rule converge to a local Nash equilibrium. arXiv: 1706.08500v6, 2017. <https://doi.org/10.48550/arXiv.1706.08500>.
15. **Shlens J.** Notes on Kullback – Leibler divergence and likelihood. arXiv: 1404.2000, 2014. <https://doi.org/10.48550/arXiv.1404.2000>.

REFERENCES

1. **Cremonesi M., Bellini C., Bian B., et al.,** Using big data technologies for HEP analysis. arXiv: 1901.07143, 2019. <https://doi.org/10.48550/arXiv.1901.07143>.
2. **Jordan M. I., Mitchell T. M.,** Machine learning: Trends, perspectives, and prospects, *Science*. 349 (6245) (2015) 255–260.
3. **Weinzierl S.,** Introduction to Monte Carlo methods. arXiv: 0006269, 2000. <https://doi.org/10.48550/arXiv.hep-ph/0006269>.
4. **Lobanov A. A., Berdnikov Ya. A.,** A generator of deep inelastic lepton-proton scattering based on the Generative-Adversarial Network (GAN), *St. Petersburg State Polytechnical University Journal. Physics and Mathematics*. 16 (4) (2023) 181–188. (in Russian).

5. **Aschenauer E. C., Borsa I., Sassot R., Van Hulse C.**, Semi-inclusive deep-inelastic scattering, parton distributions, and fragmentation functions at a future electron-ion collider, *Phys. Rev. D.* 99 (9) (2019) 094004.
6. **Barone V., Boglione M., Hernandez J. O. G., Melis S.**, Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering, *Phys. Rev. D.* 91 (7) (2015) 074019.
7. **Sjöstrand T., Mrenna S., Skands P.**, A brief introduction to PYTHIA 8.1, *Comput. Phys. Commun.* 178 (11) (2008) 852–867.
8. **Mao X., Li Q., Xie H., et al.**, On the effectiveness of least squares Generative Adversarial Networks, *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (12) (2019) 2947–2960.
9. **Sharma O.**, A new activation function for deep neural network, *Proc. Int. Conf. Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, IEEE, Faridabad, India, Febr. 14–16 (2019) 84–86.
10. **Hawkins D. M.**, The problem of overfitting, *J. Chem. Inf. Comput. Sci.* 44 (1) (2003) 1–12.
11. **Srivastava N., Hinton G., Krizhevsky A., et al.**, Dropout: A simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* 15 (2014) 1929–1958.
12. **Miyato T., Kataoka T., Koyama M., Yoshida Y.**, Spectral normalization for Generative Adversarial Networks, arXiv: 1802.05957/v1, 2018. <https://doi.org/10.48550/arXiv.1802.05957>.
13. **Xu D., Zhang Sh., Zhang H., Mandic D. P.**, Convergence of the RMSProp deep learning method with penalty for nonconvex optimization, *Neural Netw.* 139 (July) (2021) 17–23.
14. **Heusel M., Ramsauer H., Unterthiner T., et al.**, GANs trained by a two time-scale update rule converge to a local Nash equilibrium, arXiv: 1706.08500v6, 2017. <https://doi.org/10.48550/arXiv.1706.08500>.
15. **Shlens J.**, Notes on Kullback–Leibler divergence and likelihood, arXiv: 1404.2000, 2014. <https://doi.org/10.48550/arXiv.1404.2000>.

СВЕДЕНИЯ ОБ АВТОРАХ

ЛОБАНОВ Андрей Александрович – студент Физико-механического института Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
lobanov2.aa@edu.spbstu.ru
ORCID: 0000-0002-8910-4775

БЕРДНИКОВ Ярослав Александрович – доктор физико-математических наук, профессор Высшей школы фундаментальных физических исследований Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
berdnikov@spbstu.ru
ORCID: 0000-0003-0309-5917

THE AUTHORS**LOBANOV Andrey A.**

Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
lobanov2.aa@edu.spbstu.ru
ORCID: 0000-0002-8910-4775

BERDNIKOV Yaroslav A.

Peter the Great St. Petersburg Polytechnic University
29 Politechnicheskaya St., St. Petersburg, 195251, Russia
berdnikov@spbstu.ru
ORCID: 0000-0003-0309-5917

*Статья поступила в редакцию 28.09.2023. Одобрена после рецензирования 12.10.2023.
Принята 12.10.2023.*

Received 28.09.2023. Approved after reviewing 12.10.2023. Accepted 12.10.2023.