

## ОТЗЫВ

официального оппонента Деркача Д. А.

на диссертационную работу Петрова Никиты Александровича «Исследование состава космических лучей в диапазоне энергий 1,5 – 100 ПэВ на архивных данных эксперимента KASCADE с использованием машинного обучения», представленную на соискание ученой степени кандидата физико-математических наук по специальности 1.3.15 – Физика атомных ядер и элементарных частиц, физика высоких энергий.

Работа Петрова Никиты Александровича выполнена в Федеральном государственном бюджетном учреждении науки Институте ядерных исследований Российской академии наук (ИЯИ РАН). Диссертационная работа посвящена актуальной задаче физики космических лучей – исследованию энергетического спектра и массового состава космических лучей в диапазоне энергий 1,5 – 100 ПэВ. В основе работы лежит повторный анализ архивных данных эксперимента KASCADE с применением современных методов машинного обучения.

**Актуальность** темы исследования несомненна и обусловлена тем, что изучение энергетического спектра и массового состава космических лучей предоставляет ключевую информацию для понимания фундаментальных вопросов об их источниках и механизмах ускорения. Особенности спектра, такие как «коллено» в области  $\sim 4$  ПэВ, требуют детального изучения, поскольку различные теоретические модели предсказывают разное положение этой особенности и химический состав в данной области. Применение новых методов анализа, включая машинное обучение, к существующим высококачественным данным, таким как архив эксперимента KASCADE, позволяет повысить точность измерений и получить новые сведения.

Диссертация состоит из введения, трех глав, заключения, списка сокращений, списка литературы и приложения. Объем диссертации составляет 99 страниц и включает 38 рисунков, 10 таблиц, а также список литературы из 88 наименований.

**Во введении** обосновывается актуальность выбранной темы, формулируются цели и задачи исследования, отмечается научная новизна и практическая значимость полученных результатов, указаны сведения об апробации работы и публикациях по теме диссертации.

**В первой главе** приведено описание эксперимента KASCADE, используемых архивных данных и данных Монте-Карло моделирования. Описаны критерии отбора событий и используемые наборы данных Монте-Карло, сгенерированные с использованием различных моделей адронных взаимодействий (QGSJet-II.04, EPOS-LHC, Sibyll 2.3c).

**Во второй главе** подробно излагается разработанная автором методика. Описывается разработка и выбор классификатора типа первичной частицы на основе моделей машинного обучения (случайный лес, многослойный перцептрон, свёрточная нейронная сеть – CNN), где CNN показала наилучшие результаты. Проведено детальное исследование свойств CNN, включая исследование важности отдельных входных параметров, зависимость качества классификации от энергии и зенитного угла, а также влияние неработающих детекторных станций и различных генераторов адронных взаимодействий. Описана процедура анфолдинга по типу первичной частицы и по энергии. Представлен анализ систематических неопределённостей и сравнение результатов, полученных на открытой части данных (20%) с использованием генератора QGSJet-II.02, с оригинальными результатами KASCADE.

**В третьей главе** представлены результаты, полученные на закрытой части экспериментальных данных (80%) с использованием современных генераторов адронных взаимодействий (QGSJet-II.04 как основной, EPOS-LHC и Sibyll 2.3c для оценки теоретических неопределённостей). Продемонстрированы энергетические спектры пяти индивидуальных массовых компонент, проведено сравнение среднего логарифма массового числа  $\langle \ln A \rangle$  с результатами других экспериментов (IceTop, TALE, LHAASO), где показано согласие в рамках теоретических неопределённостей. Осуществлён поиск колленоподобных структур и изломов в спектрах отдельных компонент путем аппроксимации функций степенными законами (простым и ломаным), обнаружены изломы с высоким уровнем

статистической значимости в спектрах протонов и гелия, а также указание на излом в спектре железа.

**В заключении** отражены основные научные результаты, полученные в ходе выполнения диссертационной работы. **В приложении** представлены таблицы с численными значениями результатов данного исследования.

**Научная новизна** работы заключается в первом применении методики на основе машинного обучения для восстановления спектров пяти компонент по данным эксперимента KASCADE, достижении высокой точности определения энергетических спектров.

**Теоретическая значимость** работы состоит в предоставлении уточненных данных по спектрам массовых компонент, необходимых для проверки и развития моделей происхождения и ускорения галактических космических лучей. **Практическая ценность** заключается в демонстрации эффективности методов машинного обучения для анализа данных в физике космических лучей, а также в подчеркивании важности сохранения и повторного анализа данных прошлых экспериментов.

**Достоверность** полученных результатов обеспечивается тщательной проработкой методики, включая валидацию классификатора и оценку систематических неопределённостей, сравнением с результатами других экспериментов, а также апробацией на международных конференциях и публикациями в рецензируемых научных журналах.

**Личный вклад** автора является определяющим на всех этапах исследования: от предобработки данных и обучения нейронной сети до проведения процедуры анфолдинга, оценки систематических неопределённостей, анализа спектров и интерпретации результатов.

Работу отличает комплексный подход к построению структуры анализа экспериментальных данных и внимание к соответствующим деталям на всех этапах. Исследование имеет двойную ценность: с одной стороны, а наибольшей на текущий момент статистической значимостью подтверждена коленоподобная структура протонной и гелиевой массовых компонент, а также впервые получено указание на излом в железном спектре в исследуемом диапазоне энергий, с другой стороны — проведён подробный анализ подходов к компьютерному анализу данных, который методологически будет полезен будущим исследованиям.

Работа обладает несомненной значимостью с точки зрения физических наук, как указано выше, получен наиболее статистически значимый результат. Кроме того, было проверено влияние разных факторов на предложенные решения, что довольно редко делается в таком объёме. Особо следует отметить, что был проведён повторный анализ открытых данных, что демонстрирует важность сохранения и публикации данных и других экспериментов.

Можно сформулировать несколько **некритичных комментариев** по тексту:

1. При первичном анализе данных были применены критерии отбора, рекомендованные KASCADE, при этом не обсуждается влияния этих критериев, с учётом новых подходов к анализу данных.

2. Найденные коленоподобные структуры (в частности, железное) могли бы быть более подробно интерпретированы.

Указанные замечания носят дискуссионный характер и не снижают высокой научной ценности диссертационной работы. Автореферат в полной мере отражает содержание диссертации.

Диссертационная работа полностью отвечает требованиям Положения о присуждении учёных степеней, утверждённого Постановлением Правительства РФ от 24 сентября 2013 г. №842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а Петров Никита Александрович, заслуживает присуждения ему степени кандидата физико-математических наук по специальности 1.3.15 Физика атомных ядер и элементарных частиц, физика высоких энергий.

Официальный оппонент:

**Деркач Денис Александрович**, PhD, Заведующий лабораторией методов анализа больших данных факультета компьютерных наук Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ).

почтовый адрес: 109028, г. Москва, Покровский бульвар, д. 11

e-mail: dderkach@hse.ru

рабочий телефон: +7 (495) 772-95-90 27266

«02» июня 2025 г.

\_\_\_\_\_ / Д. А. Деркач

Подпись Д. А. Деркача заверяю

Специалист по персоналу

Шостак И.С.

Список основных публикаций по теме рецензируемой диссертации в рецензируемых научных изданиях за последние 5 лет:

1. D Derkach et al. RatanSunPy: A robust preprocessing pipeline for RATAN-600 solar radio observations data // *Astronomy and Computing*, 2025. – № 5. - p. 100918.
2. V. Abazovet, D. Derkach et al. (SPD Collaboration). Technical Design Report of the Spin Physics Detector at NICA // *Natural Science Review*, 2024. – № 1. - p. 1.
3. D. Derkach et al., Symbolic expression generation via variational auto-encoder // *PeerJ Computer Science*, 2023. - № 9:e1241. – p. 1.
4. D. Derkach et al., New UTfit Analysis of the Unitarity Triangle in the Cabibbo-Kobayashi-Maskawa scheme // *Rendiconti Lincei Scienze Fisiche e Naturali*, 2023. – №34. – p. 37.
5. D. Derkach et al., Stokes Inversion Techniques with Neural Networks: Analysis of Uncertainty in Parameter Estimation // *Solar Physics*, 2023. – №298. – p. 98.