

Федеральное государственное бюджетное учреждение науки
Институт ядерных исследований Российской академии наук

На правах рукописи

Петров Никита Александрович

**Исследование состава космических лучей в диапазоне энергий
1,5 – 100 ПэВ на архивных данных эксперимента KASCADE
с использованием машинного обучения**

1.3.15. Физика атомных ядер и элементарных частиц, физика высоких энергий

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
кандидат физико-математических наук
Кузнецов Михаил Юрьевич

Москва — 2025

Оглавление

	Стр.
Введение	4
Глава 1. Эксперимент KASCADE	11
1.1 KASCADE Array	11
1.2 Экспериментальные данные и Монте-Карло	13
Глава 2. Процедура восстановления спектров массовых компонент космических лучей энергии 1–100 ПэВ	19
2.1 Разработка и обучение классификаторов типа первичной частицы	19
2.2 Оценка качества реконструкции типа первичной частицы	23
2.2.1 Матрицы смешивания	23
2.2.2 Метод случайных смесей	25
2.3 Исследование классификатора типа первичной частицы	27
2.3.1 Исследование важности отдельных параметров	27
2.3.2 Зависимость от реконструированной энергии и зенитного угла первичной частицы	31
2.3.3 Неработающие детекторные станции	35
2.3.4 Зависимость от генераторов адронных взаимодействий	37
2.3.5 Спектры индивидуальных массовых компонент	40
2.4 Анфолдинг	42
2.5 Изучение систематических погрешностей	46
2.5.1 Неопределённости анфолдинга	46
2.5.2 Неопределённости Монте-Карло	50
2.5.3 Сравнительный анализ неопределённостей	52
2.6 Сравнение с оригинальным результатом KASCADE с генератором QGSJet-II.02	55
Глава 3. Результаты восстановления спектров массовых компонент космических лучей	59
3.1 Восстановление спектров массовых компонент с генератором QGSJet-II.04	59
3.2 Учёт неопределённостей, связанных с различными генераторами адронных взаимодействий	62

	Стр.
3.3 Представление спектров в виде среднего логарифма массы	67
3.4 Сравнение с результатами других экспериментов	68
3.5 Поиск коленоподобных структур в спектрах индивидуальных компонент	72
Заключение	76
Список сокращений и условных обозначений	78
Список литературы	80
Приложение А. Численные значения результатов	92

Введение

Актуальность темы

Космические лучи были открыты В. Гессом в 1912 г. [1] в серии экспериментов на аэростатах по измерению ионизации в зависимости от высоты. В результате Гесс понял, что измеряемое им излучение связано не с землёй, а с космосом. Спустя более чем сто лет мы знаем, что космические лучи состоят по большей части из высокоэнергетичных атомных ядер и элементарных частиц, однако остаётся большое количество фундаментальных вопросов связанных с ними, например, вопросы об источниках космических лучей и механизмах их ускорения [2—6]. Исследование энергетического спектра космических лучей и его массового состава может дать важную информацию для ответа на эти вопросы.

Энергетический спектр космических лучей, охватывающий диапазон энергий от ~ 10 ГэВ до 10^{11} ГэВ, в первом приближении может быть описан простым степенным законом, однако при более внимательном рассмотрении в нём на сегодняшний день обнаруживается ряд особенностей [3], таких, как отклонение от простого степенного закона для космических лучей с магнитной жёсткостью около 300 ГВ; так называемое «колени» в области $\sim 4 \cdot 10^6$ ГэВ; второе «колени» в области $\sim 3,2 \cdot 10^8$ ГэВ; «лодыжка» $10^{9,7}$ ГэВ и, наконец, обрезание спектра при $\sim 4 \cdot 10^{10}$ ГэВ. Причины возникновения этих особенностей являются предметом дискуссии, в которой одно из значимых мест занимает вопрос об интерпретации колена.

В области колена ~ 4 ПэВ энергетический спектр космических лучей резко меняет наклон, эта особенность спектра была впервые обнаружена в 1958 г. в эксперименте ШАЛ МГУ [7] и впоследствии подтверждалась в работах [8—14]. Разрабатываемые теории [15; 16] об источниках и механизмах ускорения космических лучей в качестве наблюдаемого эффекта предполагают различные позиции колена, его форму, химический состав. Часть теорий связывает колени в энергетическом спектре с максимальной энергией, которую могут достичь космические лучи в источниках [17—19], другие объясняют происхождение колена высвобождением космических лучей из нашей галактики [20]. В обоих этих случаях положение излома в спектрах индивидуальных массовых компонент прямо зависит от заряда первичной частицы. В то же время существуют более экзотичные тео-

рии [21], предполагающие новую физику и зависимость энергии колена от массы первичной частицы.

Степень разработанности темы

Впервые систематическое изменение массового состава в области колена с увеличением энергии наблюдалось в эксперименте KASCADE [13], также массовый состав изучался в работах TibetAS γ [22], KASCADE-Grande [23], IceTop [24], TALE [25], MAKET-ANI [26], Tunka-133 [27], TAIGA [28], а также недавно опубликованном анализе LHAASO [29]. И хотя другие эксперименты в целом подтверждают выводы KASCADE [13], между ними существуют разногласия, поэтому необходимо продолжать эксперименты и улучшать точность измерений.

На сегодняшний день прямое изучение космических лучей с энергиями порядка и выше ПэВ из-за их низкой интенсивности оказывается недоступным [30], и исследования проводятся только посредством регистрации вторичных частиц широких атмосферных ливней (ШАЛов) [31]. ШАЛы были открыты П. Оже в 1939 г. [32] и представляют собой каскады элементарных частиц и атомных ядер, образующихся в результате взаимодействия первичного космического луча с атомами атмосферы Земли. В таком эксперименте наземные детекторные станции регистрируют различные компоненты ШАЛа и из этих данных делаются выводы о первичной частице. В области энергий 1–100 ПэВ продолжают работать эксперименты TALE [33], Tunka [27], IceTop [34], LHAASO [35], Ковер [36].

Основная сложность такого рода экспериментов связана с тем, что требуется точное моделирование развития ШАЛа методами Монте-Карло. Для такого моделирования необходимо детальное знание электромагнитных и адронных взаимодействий, однако в случае генераторов адронных взаимодействий при высоких энергиях, которые ответственны за начальное развитие ливня, продолжает сохраняться расхождение [37] в наблюдаемых величинах для различных независимых генераторов. В настоящее время активно развиваются генераторы QGSJet [38], EPOS [39] и Sibyll [40], и ситуация в целом улучшается, поэтому важно не только анализировать новые данные экспериментальных установок, но и проводить переобработку архивных измерений с учётом новых версий генераторов.

Для повторного анализа массового состава при энергиях 1–100 ПэВ хорошо подходит эксперимент KASCADE [41], поскольку за время работы с 1996 г. по 2013 г. им был собран большой объём качественных экспериментальных дан-

ных, который впоследствии был сохранён, задокументирован и выложен в открытый доступ [42] вместе с Монте-Карло моделированием, основанным на современных версиях генераторов адронных взаимодействий. Кроме того, с момента публикации последних результатов KASCADE [23] также получили развитие методы анализа данных с использованием машинного обучения, которые могут более эффективно использовать информацию из эксперимента. Применительно к анализу массового состава космических лучей, машинное обучение использовалось ещё в эксперименте TibetASy [22], однако распространение получило сейчас [43—51] и продолжает доказывать свою полезность.

Целью данной работы является исследование массового состава космических лучей в диапазоне энергий 1,5–100 ПэВ по архивным данным эксперимента KASCADE.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. разработать методику восстановления энергетических спектров индивидуальных массовых компонент с точностью, превышающей оригинальные результаты KASCADE;
2. разработать, обучить и провалидировать классификатор, реконструирующий массовый тип первичной частицы;
3. провести процедуру анфолдинга реконструированного типа и энергии первичных частиц, учесть систематические неопределённости;
4. проанализировать поведение спектров в зависимости от энергии, проверить спектры на предмет коленоподобных структур и изломов.

Научная новизна

1. Впервые была разработана и применена к данным эксперимента KASCADE методика по восстановлению спектров пяти массовых компонент (протонной, гелиевой, углеродной, кремниевой, железной) с использованием моделей машинного обучения.
2. Определены энергетические спектры пяти массовых компонент в диапазоне 1,5–100 ПэВ с точностью превышающей исследования KASCADE и других экспериментов.
3. С 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.

Теоретическая и практическая значимость

- Независимое измерение энергетических спектров массовых компонент в области энергий 1,5–100 ПэВ необходимо для проверки теоретических моделей, объясняющих происхождение и механизмы ускорения галактических космических лучей.
- Использование методов машинного обучения для анализа данных подобных экспериментов открывает путь к увеличению точности результатов, благодаря более эффективному использованию информации.
- Повторный анализ открытых архивных экспериментальных данных с использованием новых методов демонстрирует важность сохранения и публикации данных и других экспериментов.

Методология и методы исследования

В основе диссертации лежат методы машинного обучения (нейронные сети, случайный лес) для решения задачи классификации массового типа первичной частицы, методы Монте-Карло, методы анфолдинга, а также методы статистической обработки данных.

Основные положения, выносимые на защиту

1. Разработанная методика с использованием методов машинного обучения позволяет определить энергетические спектры пяти индивидуальных массовых компонент (протоны, гелий, углерод, кремний, железо) в диапазоне энергий 1,5–100 ПэВ на архивных данных KASCADE.
2. Систематическая неопределённость энергетических спектров пяти индивидуальных массовых компонент (протоны, гелий, углерод, кремний, железо) космических лучей в диапазоне энергий 1,5–100 ПэВ на архивных данных KASCADE, полученных с помощью разработанной методики, ниже соответствующей в оригинальных работах KASCADE и в современных экспериментах (IceTop).
3. Получена оценка неопределённостей в спектрах индивидуальных массовых компонент, связанная с использованием трёх различных современных генераторов адронных взаимодействий: QGSJet-II.04, EPOS-LHC, Sibyll 2.3c.
4. Энергетический спектр протонной компоненты демонстрирует избыток (~ 10 раз при энергии 10 ПэВ), а спектры остальных компонент — недостаток относительно оригинальных результатов KASCADE.

5. Коленоподобное поведение в спектрах протонной и гелиевой массовых компонент наблюдается со значимостью более 5σ и 3σ при энергиях $\sim 4,4$ ПэВ и ~ 11 ПэВ соответственно. Указание на излом в спектре железной компоненты обнаружено при энергии $\sim 4,5$ ПэВ.

Степень достоверности и апробация результатов

Достоверность полученных результатов обеспечивается с помощью различных проверок и тестов, проведённых в данной работе, сравнением полученных результатов с другими экспериментами, а также публикацией результатов данного исследования в рецензируемых научных журналах и обсуждениями на научных конференциях.

Основные результаты работы докладывались на конференциях:

1. Workshop on Machine Learning for Cosmic-Ray Air Showers (Delaware, USA, 2022);
2. The 27th European Cosmic Ray Symposium (ECRS, Nijmegen, Netherlands, 2022);
3. International Conference on Particle Physics and Cosmology (Yerevan, Armenia, 2023);
4. 38-я Всероссийская конференция по космическим лучам (Москва, Россия, 2024).

Результаты также демонстрировались и обсуждались на семинаре лаборатории обработки больших данных ИЯИ РАН 17.01.2024, экспериментальном семинаре ИЯФ СО РАН 01.03.2024; были представлены на конкурсе молодых учёных ИЯФ СО РАН 14.05.2024.

Публикации

По теме диссертации опубликовано 3 работы, из них 2 в научных журналах из списка ВАК [52; 53] а также 1 работа в трудах конференций [54]. Все 3 публикации проиндексированы в базе данных Scopus.

Личный вклад автора

Личный вклад автора в получение научных результатов, лежащих в основе диссертации является определяющим. Автором выполнены все этапы по предварительной обработке данных и обучению свёрточной нейронной сети в качестве классификатора типа первичной частицы, разработаны и проведены необходимые проверки по валидации классификатора. Автор провёл анфолдинг и получил полный и индивидуальные энергетические спектры массовых компонент в диапазоне энергий от 1,5 ПэВ до 100 ПэВ с учётом оценки систематических неопре-

делённостей. А также, автором выполнена процедура по поиску коленоподобных структур и изломов в полученных энергетических спектрах и интерпретация результатов.

Авторский вклад в статье [52] заключается в следующем: автором проведена предварительная подготовка данных для обучения классификатора типа первичной частицы; предложена и обучена свёрточная нейронная сеть, проведены исследования по оптимизации параметров архитектуры сети; реализован метод случайных смесей для демонстрации и сравнения качества работы различных классификаторов. Автор исследовал влияние входных параметров на результат работы свёрточной нейронной сети, провёл ряд тестов с целью проверки корректности её работы, реализовал сравнение качества результатов реконструкции типа первичной частицы, полученных на Монте-Карло моделировании с использованием различных генераторов адронных взаимодействий. Автор предложил и реализовал процедуру анфолдинга энергетических спектров отдельных массовых компонент, провёл систематизацию, анализ и оценку вклада источников систематических неопределённостей, а также продемонстрировал прямое сравнение результатов с последними оригинальными результатами эксперимента KASCADE на 20 % экспериментальных данных.

В статье [53] авторский вклад состоит в следующем: автором была применена методика, описанная в статье [52], к закрытой части (оставшиеся 80 %) экспериментальных данных, проведена оценка неопределённостей, связанных с использованием различных генераторов адронных взаимодействий, реализовано представление полученных результатов в виде зависимости среднего логарифма массы от энергии первичной частицы. Автор исследовал поведение энергетических спектров отдельных компонент на предмет коленоподобных структур и изломов, сравнил полученные результаты с другими экспериментами, а также провёл интерпретацию полученных результатов.

Структура и объем диссертации

Диссертация состоит из введения, 3 глав, заключения, списка сокращений и условных обозначений, списка литературы и 1 приложения. Глава 1 посвящена описанию эксперимента KASCADE, а также используемых в данной работе Монте-Карло и экспериментальных данных. В Главе 2 описывается предложенная методология по определению энергетических спектров индивидуальных массовых компонент, в частности разработка и обучение классификатора типа первичной частицы на основе свёрточной нейронной сети, а также проце-

дура анфолдинга по типу и энергии первичной частицы. Глава 3 демонстрирует результаты применения разработанной методики к данным с использованием современных генераторов адронных взаимодействий QGSJet-II.04, EPOS-LHC и Sibyll 2.3c. А именно, получены энергетические спектры индивидуальных массовых компонент, зависимость среднего логарифма массы от энергии первичной частицы, результаты поиска коленоподобных структур и изломов в наблюдаемых спектрах, а также сравнение полученных результатов с другими экспериментами в изучаемом энергетическом диапазоне. В заключении перечисляются основные результаты исследования. В Приложении А представлены таблицы с численными значениями основных результатов данной работы.

Полный объём диссертации составляет 99 страниц, включая 38 рисунков и 10 таблиц. Список литературы содержит 88 наименований.

Глава 1. Эксперимент KASCADE

KASCADE (KArlsruhe Shower Core and Array DEtector) — это эксперимент по изучению состава первичных космических лучей и адронных взаимодействий в энергетическом диапазоне 500 ТэВ–100 ПэВ посредством регистрации широких атмосферных ливней (ШАЛов) [41]. KASCADE располагался в Карлсруэ, Германия (49,1° северной широты, 8,4° восточной долготы) на высоте 110 м над уровнем моря, соответствующей средней атмосферной глубине 1022 г/см². Эксперимент KASCADE (вместе со своим расширением KASCADE-Grande [55]) непрерывно набирал данные с ноября 1996 г. по январь 2013 г.

Эксперимент состоял из четырёх основных детекторных компонент:

- массива из 252 детекторных станций (KASCADE Array), регистрирующих отдельно электронную и мюонную компоненты;
- центрального детектора (Central Detector) для измерения адронной компоненты;
- туннеля со стримерными трубками (Muon-Tracking Detector) для записи отдельных мюонных треков;
- дополнительного массива из 37 детекторных станций (Grande Detector Array) для регистрации электронной компоненты, установленного в 2003 г.

В текущем исследовании используются данные только с KASCADE Array, поэтому далее будет представлено подробное описание именно этой компоненты эксперимента.

1.1 KASCADE Array

KASCADE Array состоял из 252 детекторных станций, расположенных в узлах прямоугольной сетки с шагом 13 м и покрывал площадь 200 × 200 м². Общая схема установки проиллюстрирована на рисунке 1.1, а на рисунке 1.2 изображены схемы детекторных станций KASCADE Array.

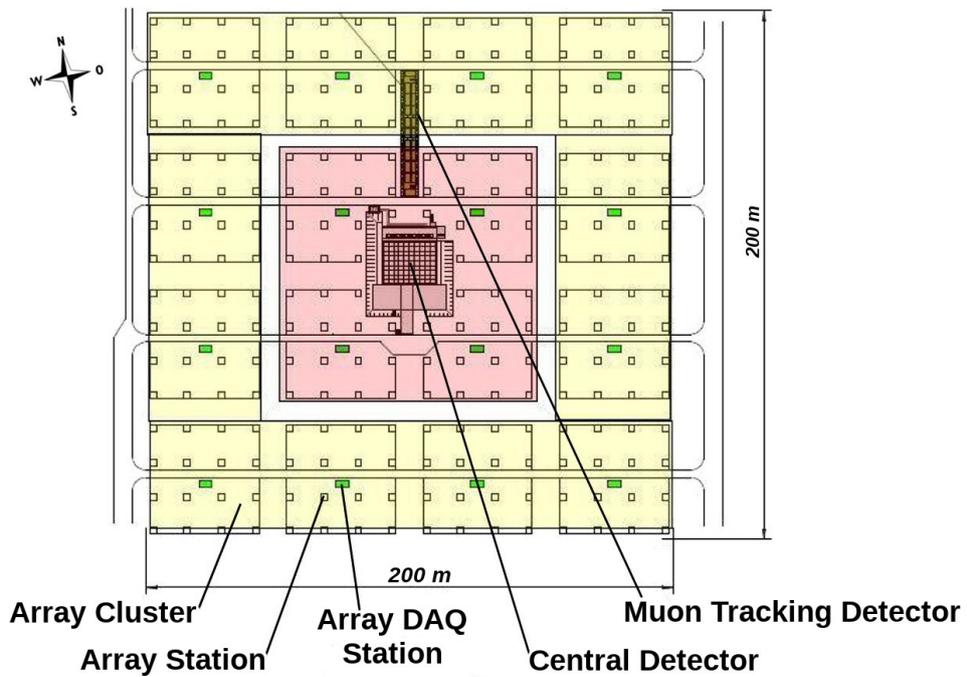


Рисунок 1.1 — Схема KASCADE Array. Зелёным выделены детекторные станции типа I, регистрирующие электромагнитную и мюонную компоненты ШАЛов. Красным выделены детекторные станции типа II, расположенные в центральной области, регистрирующие только электромагнитную компоненту ШАЛов

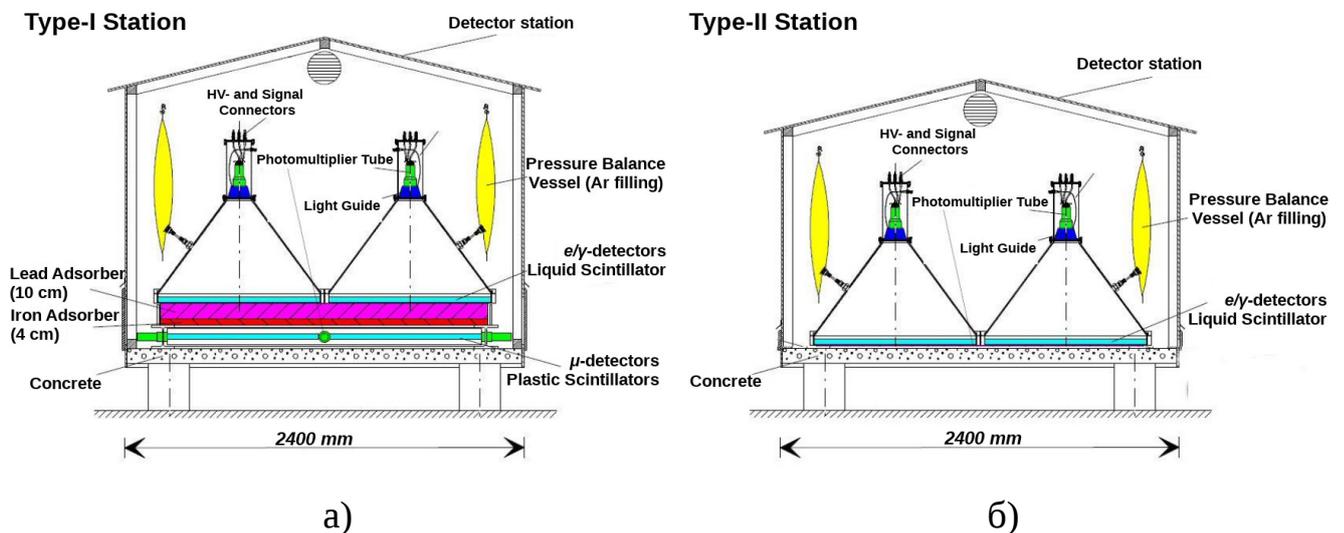


Рисунок 1.2 — Схемы детекторных станций KASCADE Array (а) типа I и (б) типа II

Как видно из общей схемы, станции представляли собой один из двух типов и объединены в кластеры размером 4×4 . Станции внутренних 4 кластеров были оборудованы слоем жидкого сцинтиллятора с эффективной площадью $3,14 \text{ м}^2$ для определения электромагнитной компоненты. Остальные станции,

кроме сцинтиллятора для определения электромагнитной компоненты, оборудованы пластиковым сцинтиллятором для выделения мюонной компоненты ливня, который находился под защитным слоем из 10 см свинца и 4 см железа.

Также каждый электромагнитный сцинтиллятор был оснащён ФЭУ для регистрации времени высвечивания. Их временное разрешение составляло 0,77 нс, что приводит к угловому разрешению порядка $0,1^\circ$ [56]. Энергетическое разрешение составляло 8 % для электромагнитной и 10 % для мюонной компоненты [41].

1.2 Экспериментальные данные и Монте-Карло

В данной работе используются экспериментальные данные и моделирование Монте-Карло эксперимента KASCADE, а именно её основной компоненты KASCADE Array, описанной выше. Все данные предоставлены проектом KASCADE Cosmic Ray Data Centre (KCDC) [42], который опубликовал в открытом доступе предобработанные экспериментальные данные и наборы данных моделирования KASCADE, а также необходимую документацию для работы с ними.

Процесс моделирования устроен следующим образом, сначала происходит симуляция ШАЛов в атмосфере, для этой цели используются программа CORSIKA, а т.к. в области высоких энергий существует большой набор различных генераторов (моделей) адронных взаимодействий, то KCDC предоставляет наборы данных, полученные с помощью современных (post-LHC) генераторов адронных взаимодействий трёх основных семейств: QGSJet-II.04 [57], EPOS-LHC [39], Sibyll 2.3c [58]. Этими генераторами описываются взаимодействия частиц с энергиями выше 200 ГэВ. Взаимодействия частиц с энергиями ниже 200 ГэВ моделируются с помощью генератора FLUKA [59]. Затем моделирование проходит через программу CRES (Cosmic Ray Event Simulation), в которой происходит симуляция отклика детектора с помощью пакета GEANT 3 [60].

Для событий моделирования и эксперимента производится реконструкция некоторых параметров ливня программным пакетом KRETA (KASCADE Reconstruction for Extensive Airshowers). Таким образом, каждое событие содержит следующий набор параметров: времена прибытия (временная метка

попадания первой заряженной частицы для каждой детекторной станции), e/γ и μ энерговыделения в детекторных станциях, а также набор основных наблюдаемых параметров ШАЛа: позиция ядра ливня (x, y) , зенитный и азимутальный углы прихода первичной частицы (θ, φ) , интегральное количество электронов и мюонов (N_e, N_μ) , а также «возраст» ливня (s) .

Параметры N_e , N_μ и s определяются с помощью аппроксимации модифицированной функцией пространственного распределения Нишимуры-Каваты-Грейзена (ФПР НКГ) электронов и мюонов следующего вида [56]:

$$\rho_{e,\mu}(r) = C(s)N_{e,\mu} \left(\frac{r}{r_m^{e,\mu}} \right)^{s-\alpha} \left(1 + \frac{r}{r_m^{e,\mu}} \right)^{s-\beta},$$

$$C(s) = \frac{\Gamma(\beta - s)}{2\pi r_m^2 \Gamma(s - \alpha + 2)(\alpha + \beta - 2s)},$$

где $\rho_{e,\mu}(r)$ — плотность электронов/мюонов как функция расстояния до ядра ливня, $r_m^{e,\mu}$ — мольеровский радиус для электронов и мюонов, оценивался из Монте-Карло и составил $r_m^e = 89$ м и $r_m^\mu = 420$ м, α и β — параметры, также извлекаемые из Монте-Карло моделирования.

Дополнительно для каждого события реконструируется энергия первичной частицы (E) по следующей формуле

$$\log_{10}(E/\text{эВ}) = 1,934\,99 + 0,257\,88 \cdot \log_{10}(N_e) + 0,667\,04 \cdot \log_{10}(N_\mu) + 0,075\,07 \cdot \log_{10}^2(N_e) + 0,092\,77 \cdot \log_{10}^2(N_\mu) - 0,161\,31 \cdot \log_{10}(N_e) \log_{10}(N_\mu) \quad (1.1)$$

Эта формула получена из аппроксимации моделирования Монте-Карло с генератором QGSJet-II.02 [61].

Примеры событий из эксперимента и моделирования Монте-Карло показаны на рисунке 1.3.

В данной работе к событиям применяются используемые в оригинальной работе KASCADE [13] критерии отбора, а именно

- $\theta < 18^\circ$;
- $\sqrt{x^2 + y^2} < 91$ м;
- $\log_{10} N_e > 4,8$;
- $\log_{10} N_\mu > 3,6$;
- $0,2 < s < 1,48$ (отбор рекомендован KCDC [56] и строже оригинального $0,2 < s < 2,1$).

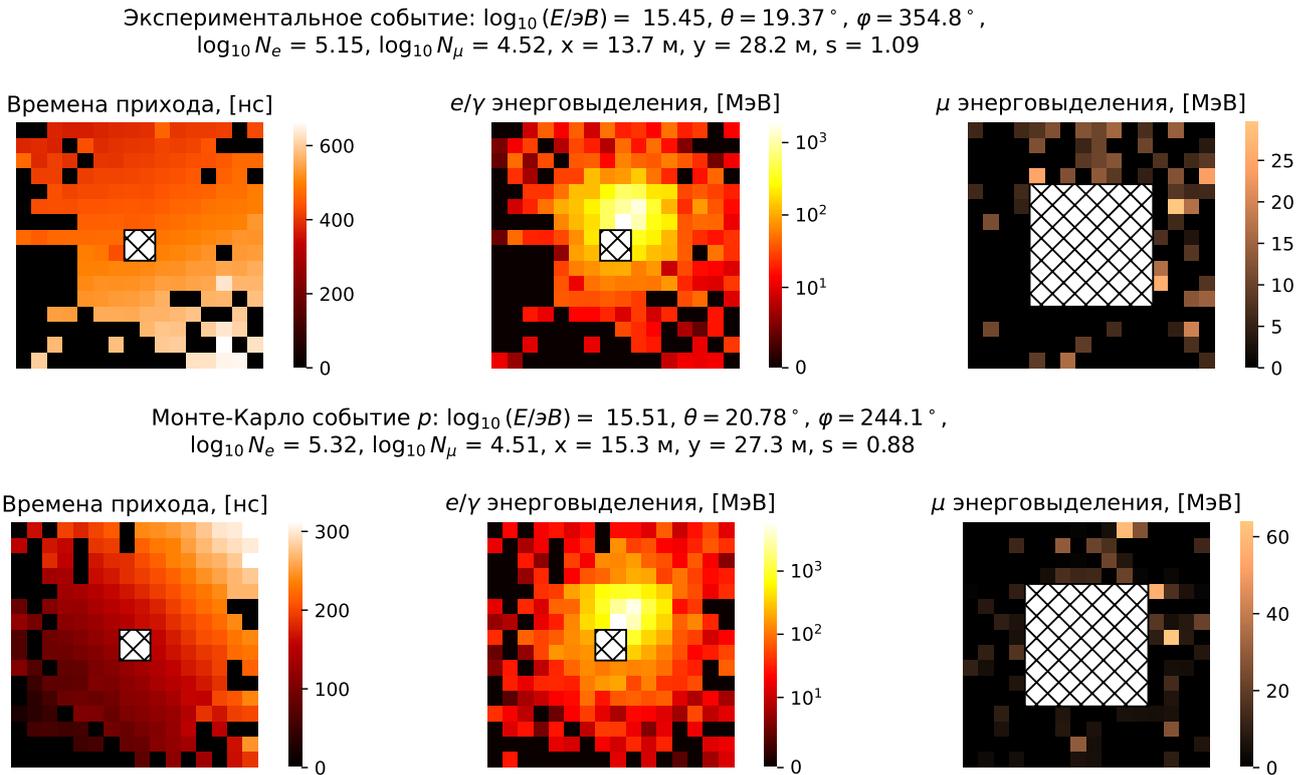


Рисунок 1.3 — Примеры событий KASCADE из эксперимента (верхняя часть) и Монте-Карло протонов (нижняя часть). Продемонстрированы реконструированные параметры каждого из событий, а также времена прихода (слева), энергосвечения электромагнитной (центр) и мюонной компоненты (справа) со станций. В KASCADE Array нет детекторных станций в центральной области 2×2 , измеряющих времена прибытия и электромагнитную компоненту, а в области 8×8 — мюонную компоненту. Эти области показаны штриховкой на рисунке

Полная эффективность триггера достигается при энергиях $E \gtrsim 10^{15}$ эВ, а эффективность используемых критериев отбора зависит от массовой компоненты и достигает 0,99 при энергии $10^{15,15}$ эВ для протонов и $10^{15,33}$ эВ для ядер железа. Эта неэффективность учитывается при измерении энергетических спектров.

В данной работе исследуется массовый состав космических лучей в энергетической области от $10^{15,15}$ эВ до 10^{17} эВ, поскольку как было отмечено выше, в этой области энергий эффективность критериев отбора начинает выходить на единицу, а также становится стабильной работа классификатора типа первичной частицы (подробнее в Главе 2). Таким образом, все последующие рисунки, иллюстрирующие качество классификации, будут приводиться в изучаемом диапазоне энергий, если не указано иное.

В итоге после применения критериев отбора получен экспериментальный набор данных объёмом $3,5 \cdot 10^6$ событий в изучаемом диапазоне энергий, который в ~ 13 раз больше набора, анализируемого в последнем исследовании данных KASCADE [23; 62]. Полная экспозиция (\mathcal{E}) после применения критериев отбора вычислялась как $\mathcal{E} = A \times T_{lt}$, где $A = (\pi \cdot \sin^2(18^\circ)) \cdot (\pi \cdot (91 \text{ м})^2)$ — апертура детектора, T_{lt} — полное время работы экспериментальной установки.

Для определения полного времени работы использовалась следующая методика [63]: регистрация космических лучей рассматривалась как пуассоновский процесс, в таком случае временные интервалы между двумя последовательно зарегистрированными событиями (Δt) распределены согласно экспоненциальному закону со следующей плотностью вероятности

$$f(t) = \frac{\exp(-\Delta t/\Delta t_0)}{\Delta t_0},$$

где Δt_0 — среднее время между последовательной регистрацией двух событий, а, в свою очередь, полное время работы

$$T_{lt} = \Delta t_0 \cdot N_{tot},$$

где N_{tot} — число зарегистрированных событий за время работы установки. Поскольку используемые экспериментальные данные содержат поле с указанием момента времени регистрации каждого события [56], то в результате удалось определить, что экспериментальный набор данных соответствует $T_{lt} = 4,2 \cdot 10^8$ с.

В данной работе экспериментальный набор данных случайным образом разделяется на две части — открытую и закрытую — в соотношении 20 % : 80 %. Открытая часть использовалась для отработки и проверки разработанной в данной работе методики анализа, а также для прямого сравнения полученных спектров массовых компонент с последними оригинальными результатами KASCADE [23] в рамках pre-LHC адронной модели QGSJet-II.02. После того как методика была разработана и проверена, она фиксировалась, и проводилось измерение спектров массовых компонент на закрытой части данных с использованием трёх наборов Монте-Карло, основанных на современных (post-LHC) моделях адронных взаимодействий как QGSJet-II.04, EPOS-LHC и Sibyll 2.3c.

Как было упомянуто выше, в данном исследовании использовались несколько наборов Монте-Карло моделирования, предоставленных KCDC, которые основаны на QGSJet-II.04 + FLUKA (ver. 2012.2.14_32), EPOS-LHC +

FLUKA (ver. 2011.2b.4_32), Sibyll 2.3c + FLUKA (ver. 2011.2c.3_64), QGSJet-II.02 + FLUKA (ver. 2002.4) моделях адронных взаимодействий и содержат $1,3 \cdot 10^5$, $6,7 \cdot 10^4$, $6,8 \cdot 10^4$, $3,7 \cdot 10^4$ событий соответственно после применения критериев отбора в изучаемом диапазоне энергий. Монте-Карло генерируется для пяти массовых групп: протонов (p), гелия (He), углерода (C), кремния (Si) и железа (Fe). Для Монте-Карло моделирования, основанного на QGSJet-II.04 адронной модели, соотношение массовых групп после применения критериев отбора в изучаемом диапазоне энергий выглядит следующим образом: p 28 %, He 22 %, C 19 %, Si 16 %, Fe 15 %. Похожее соотношение сохраняется и для других наборов Монте-Карло. Изначально в Монте-Карло генерируется одинаковое количество событий для всех типов первичных частиц, приведённое выше соотношение является следствием неединичной эффективности критериев отбора на нижнем крае исследуемого диапазона энергий. Энергетический спектр событий в Монте-Карло установлен как $dN/dE \sim E^{-2}$ и перевзвешивался к $E^{-2.7}$ в ходе анализа.

Дополнительно стоит отметить, что энергетическое разрешение как десятичный логарифм отношения реконструированной энергии E_{rec} по формуле (1.1) к истинной энергии первичной частицы E_{true} , заложенной в симуляции, для событий Монте-Карло, прошедших используемые критерии отбора в изучаемом диапазоне энергий составило около 10 % в случае генератора QGSJet-II.02. Распределение числа событий по $\log_{10}(E_{rec}/E_{true})$ для этого случая показано на Рисунке 1.4. В случае генератора QGSJet-II.04 разрешение незначительно выше и составило около 11 %.

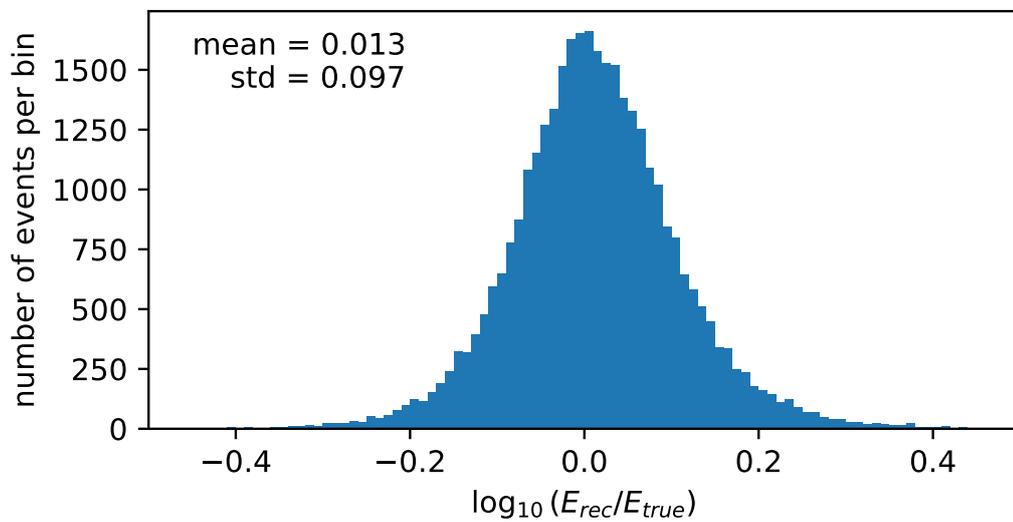


Рисунок 1.4 — Распределение числа событий в Монте-Карло моделировании с генератором QGSJet-II.02, прошедших используемые критерии отбора, по десятичному логарифму отношения между реконструированной энергией (E_{rec} по формуле (1.1)) и истинной энергией первичной частицы (E_{true}). Также представлены выборочное среднее (mean), и стандартное отклонение (std)

Глава 2. Процедура восстановления спектров массовых компонент космических лучей энергии 1–100 ПэВ

В оригинальных работах KASCADE [13; 23; 62] по восстановлению спектров массовых компонент используется анфолдинг (unfolding) [64], применяемый к распределению экспериментальных событий по реконструированным переменным $\log_{10} N_e$, $\log_{10} N_\mu$, с целью получить распределение событий по типу и энергии первичной частицы. Анфолдинг там производится с помощью матриц, определяющих соответствие между двумя этими распределениями, извлечённых из моделирования Монте-Карло. Несмотря на кажущуюся простоту, метод требует особой внимательности, поскольку прямое разделение типа массовых компонент по $\log_{10} N_e$, $\log_{10} N_\mu$ слабое и анфолдинг может оказаться нестабильным; также для двумерного анфолдинга необходимо большое количество моделирования Монте-Карло, чтобы установить распределение по $\log_{10} N_e$, $\log_{10} N_\mu$ для каждого бина по энергии и типу первичной частицы.

В данной работе приводится следующий метод восстановления спектра массовых компонент: на первом этапе реконструируется тип первичной частицы, для этой цели был разработан и обучен соответствующий классификатор с использованием машинного обучения; на втором этапе проводится последовательный анфолдинг распределения событий по реконструированной энергии первичной частицы и, затем, по реконструированному типу первичной частицы. Такой подход позволяет использовать большее количество информации о событии в сравнении с оригинальным методом, в том числе информацию об энерговыделении с отдельных станций KASCADE Array, и, как следствие, повысить качество реконструкции отдельных массовых компонент и сделать последующий анфолдинг более стабильным.

2.1 Разработка и обучение классификаторов типа первичной частицы

Был разработан и обучен ряд классификаторов с использованием машинного обучения для пособытийного определения типа первичной частицы. В качестве самой простой базовой модели был выбран классификатор на основе метода

случайного леса (Random Forest classifier — RF) [65; 66], эта модель использовала в качестве входных аргументов реконструированные параметры ШАЛ. Далее были разработаны и обучены нейронные сети, в частности, свёрточная нейронная сеть (convolutional neural network — CNN), многослойный перцептрон (multilayer perceptron — MLP), а также EfficientNet v2 [67]. Эти модели уже использовали не только реконструированные параметры, но и энерговыделения со станций KASCADE Array. Все модели классифицировали каждое событие по одному из пяти массовых типов: p , He , C , Si , Fe и были реализованы на языке программирования Python.

Для каждого метода было обучено несколько моделей на моделировании Монте-Карло из различных семейств генераторов адронных взаимодействий: QGSJet-II.04, EPOS-LHC, Sibyll 2.3c, а также на Монте-Карло QGSJet-II.02 для прямого сравнения результатов с последним оригинальным анализом KASCADE. Наборы Монте-Карло моделирования делились на тренировочную, валидационную и тестовую части в пропорции 70 % : 20 % : 10 % соответственно; тренировочные и валидационные части использовались для обучения моделей, а тестовые части — для измерения качества классификации.

На вход в RF подавались все доступные реконструированные параметры, а именно E , x , y , θ , φ , $\log_{10} N_e$, $\log_{10} N_\mu$, s . Гиперпараметры RF (число деревьев, максимальная глубина дерева, критерий качества разделения) подбирались с помощью алгоритма Grid Search для максимизации точности классификации. Программная реализация модели выполнена с использованием пакета Scikit-learn [68].

Следующий разработанный в данном исследовании классификатор — это многослойный перцептрон — нейронная сеть, состоящая из двух скрытых линейных слоёв, на каждом из которых установлена пакетная (batch) нормализация [69], ELU активация [70] и дропаут [71] (на уровне 0,15). Классификатор обучался с использованием оптимизатора Adam [72] на пакетах размером 1024. На вход модели подавались e/γ и μ энерговыделения с детекторных станций KASCADE Array, а также реконструированные параметры θ , φ . Получившаяся MLP содержала $\sim 200\,000$ обучаемых параметров, программная реализация была выполнена с использованием пакета TensorFlow [73].

Затем была разработана и обучена свёрточная нейронная сеть, архитектура которой представлена на Рисунке 2.1. Эта CNN вдохновлена классической архитектурой LeNet-5 [74] и состоит из двух основных частей: первая часть содержит

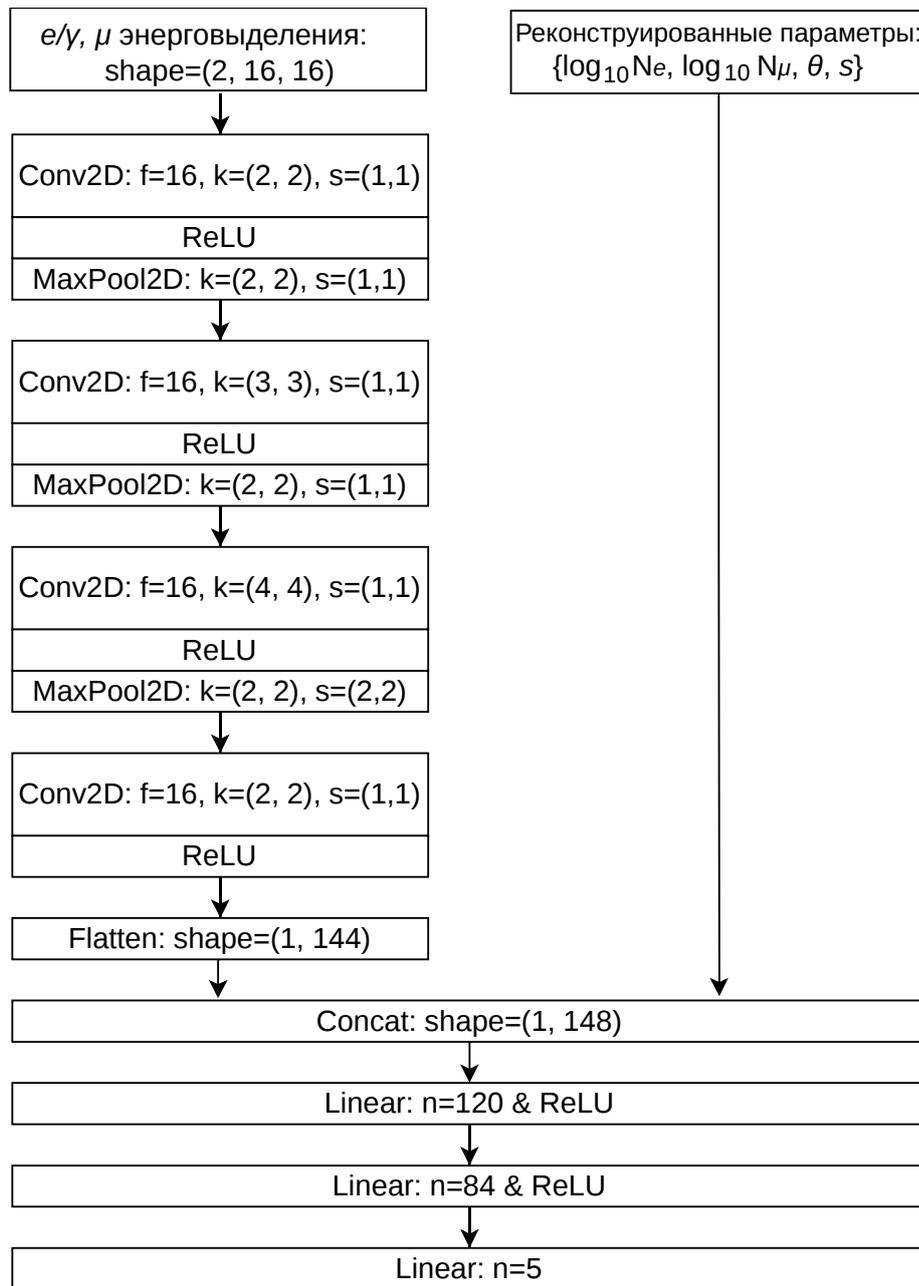


Рисунок 2.1 — Архитектура CNN, используемой в данной работе. Движение данных показано стрелками. Наименования слоёв соответствуют используемым в PyTorch. Сокращения: «f» — количество фильтров, «k» — размер ядра свёртки (kernel size), «s» — шаг фильтра свёртки (stride), «n» — количество нейронов в слое, «shape» — форма данных

набор свёрточных фильтров, которые применяются к e/γ и μ энерговыведениям с детекторных станций KASCADE Array, что позволяет извлечь паттерны, характерные для различных массовых компонент; вторая часть состоит из нескольких линейных слоёв с ReLU активацией [75], на которые подаются результаты из первой части вместе с реконструированными параметрами ливня, и комбинируются для максимизации точности классификации отдельных компонент. Данная свёр-

точная сеть имела $\sim 30\,000$ обучаемых параметров, программная реализация была выполнена с помощью фреймворка PyTorch [76]. Процесс обучения происходил с использованием оптимизатора Adam на пакетах размером 512 и контролировался в программе Weights&Biases [77].

Таким образом, на вход CNN подаются e/γ и μ энерговыделения в виде двухканального изображения размера 16×16 и реконструированные параметры $\log_{10} N_e$, $\log_{10} N_\mu$, s , θ . Входные аргументы были предварительно обработаны: энерговыделения нормализованы, т. е. приведены к нулевому среднему и единичному стандартному отклонению; реконструированные параметры перенормированы в диапазон от 0 до 1 (все необходимые коэффициенты для пересчёта получены из Монте-Карло моделирования протонов с генератором QGSJet-II.04). Дополнительно применялись аугментации, например, вращение изображения с энерговыделениями на углы кратные 90° , для увеличения стабильности тренировочного процесса. Также в обучении этой сети задействованы данные с более слабым ограничением на зенитный угол, а именно $\theta < 30^\circ$, и без ограничений на реконструированную энергию первичной частицы, поскольку в этом диапазоне данные ещё остаются качественными, о чём пишет KCDC [56], а увеличение объёма тренировочного набора данных только положительно сказывается на стабильности и качестве результатов классификатора.

Наконец, последней разработанной и обученной моделью машинного обучения стала модель, основанная на стандартной архитектуре EfficientNetV2-S. Это, так же как и предыдущая CNN модель, свёрточная архитектура, но с дополнительными оптимизациями и значительно большим числом параметров, а именно ~ 22 миллиона; её использование необходимо как источник дополнительной проверки результатов CNN, поскольку EfficientNetV2-S является хорошо зарекомендовавшей себя архитектурой, часто используемой в задачах, связанных с машинным обучением. Эта нейронная сеть принимала на вход свёрточных слоёв e/γ , μ энерговыделения со станций, а также реконструированные параметры θ и φ на внутреннем слое сети. Так же как и предыдущие сети, EfficientNetV2-S обучалась с помощью оптимизатора Adam на пакетах размером 128. По результатам обучения, было отмечено, что качество классификации этой моделью практически идентично результатам, полученным на CNN, однако поскольку EfficientNetV2-S архитектура значительно более комплексная, это говорит о том, что увеличение сложности свёрточной сети не приводит к какому-либо значительному улучшению качества итоговых результатов, а также, что резуль-

таты разработанной в данном исследовании CNN достаточны для репрезентации с её помощью всего класса свёрточных сетей. Таким образом, эта модель, более ресурсоёмкая и тяжёлая в обучении, сложна для практического использования и в дальнейшем не будет обсуждаться.

2.2 Оценка качества реконструкции типа первичной частицы

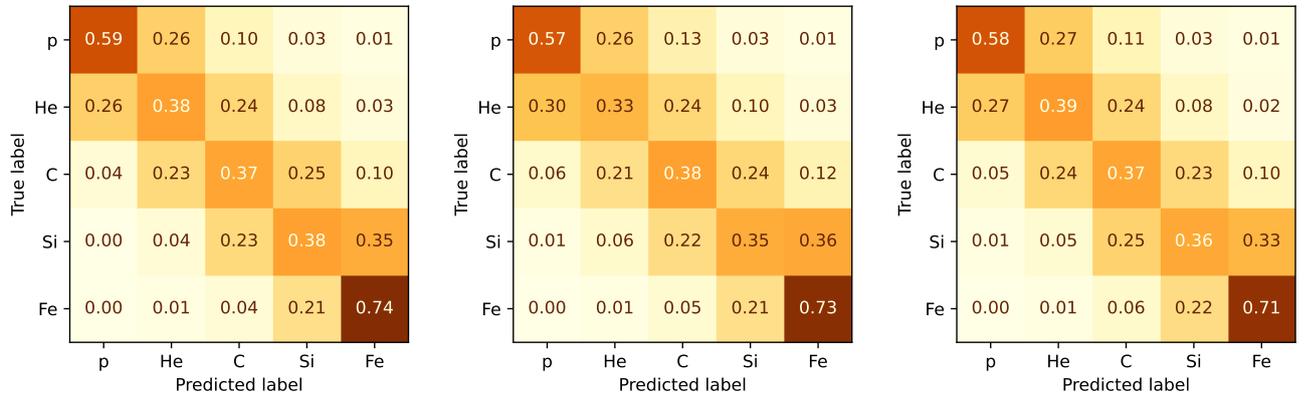
Для сравнения обученных классификаторов друг с другом и выбора лучшего среди них необходимо разработать процедуру оценки качества реконструкции типа первичной частицы соответствующим классификатором.

В данном разделе приводится определение матриц смешивания, а также метода случайных смесей, с помощью которых определяется основной классификатор, использующийся далее в исследовании.

2.2.1 Матрицы смешивания

Для оценки качества реконструкции типа первичной частицы моделями машинного обучения, описанными выше, в данной работе будут использоваться матрицы смешивания (confusion matrices). Каждая строка такой матрицы представляет собой набор вероятностей отнести частицу к реконструированному типу для соответствующего истинного типа, заданного в Монте-Карло моделировании. Или, в терминах условных вероятностей, $C_{ij} = P(A_j^{\mathcal{R}} | A_i^{\mathcal{T}})$, где \mathbf{C} — матрица смешивания, $A_i^{\mathcal{R}/\mathcal{T}}$ — реконструированный / истинный тип первичной частицы (т. е. вектор A_i состоит из типов частиц: $\{p, He, C, Si, Fe\}$). Основное свойство этой матрицы состоит в том, что чем более диагональна матрица смешивания и чем ближе к единичной матрице, тем выше качество классификации; соответственно, единичная матрица является демонстрацией результатов идеального классификатора; а в случае случайного угадывания для текущей задачи разделения событий по пяти массовым типам, во всех ячейках матрицы смешивания располагались бы значения 0,2.

На рисунках 2.2 — 2.4 показаны матрицы смешивания для RF, MLP и CNN моделей машинного обучения, обученных на Монте-Карло наборах данных с генераторами QGSJet-II.04, EPOS-LHC и Sibyll 2.3с соответственно.

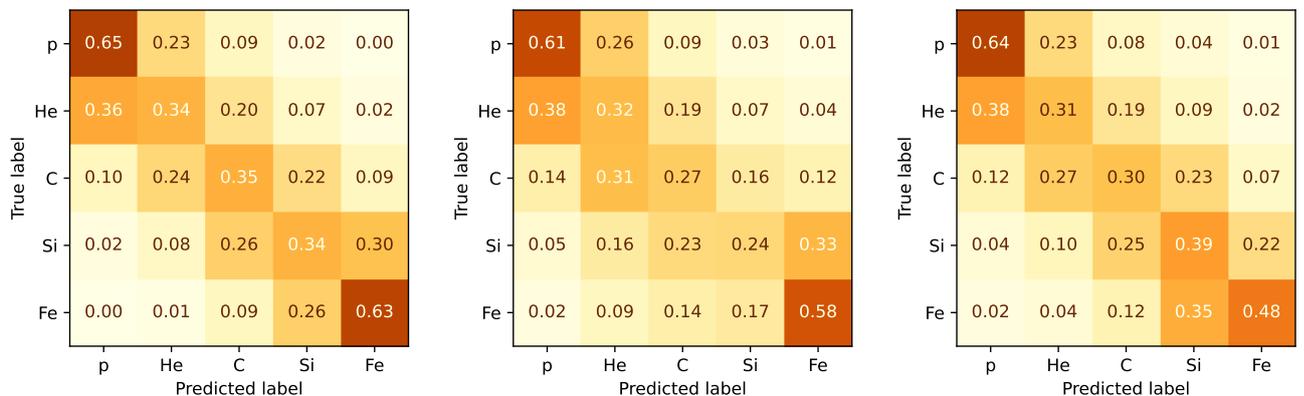


а)

б)

в)

Рисунок 2.2 — Матрицы смешивания для моделей RF обученных и протестированных на трёх наборах моделирования Монте-Карло с генераторами (а) QGSJet-II.04, (б) EPOS-LHC, (в) Sibyll 2.3с соответственно



а)

б)

в)

Рисунок 2.3 — Матрицы смешивания для моделей MLP обученных и протестированных на трёх наборах моделирования Монте-Карло с генераторами (а) QGSJet-II.04, (б) EPOS-LHC, (в) Sibyll 2.3с соответственно

Матрицами продемонстрировано качество классификации на тестовых частях Монте-Карло моделирования с учётом критериев отбора, описанных в параграфе 1.2, в изучаемом диапазоне энергий. Важно отметить несколько пунктов: качество реконструкции типа первичной частицы для всех обученных классификаторов намного превосходит случайное угадывание; для большинства моделей наибольшие значения вероятностей достигаются в диагональных элементах;

возможных вариантов пропорций и поэтому хорош для оценки качества работы классификатора. Пример полученных распределений по абсолютным отклонениям предсказанной доли первичной частицы от заложенной в смеси для CNN классификатора, обученного и протестированного на Монте-Карло моделировании с генератором QGSJet-II.04 показано на Рисунке 2.5.

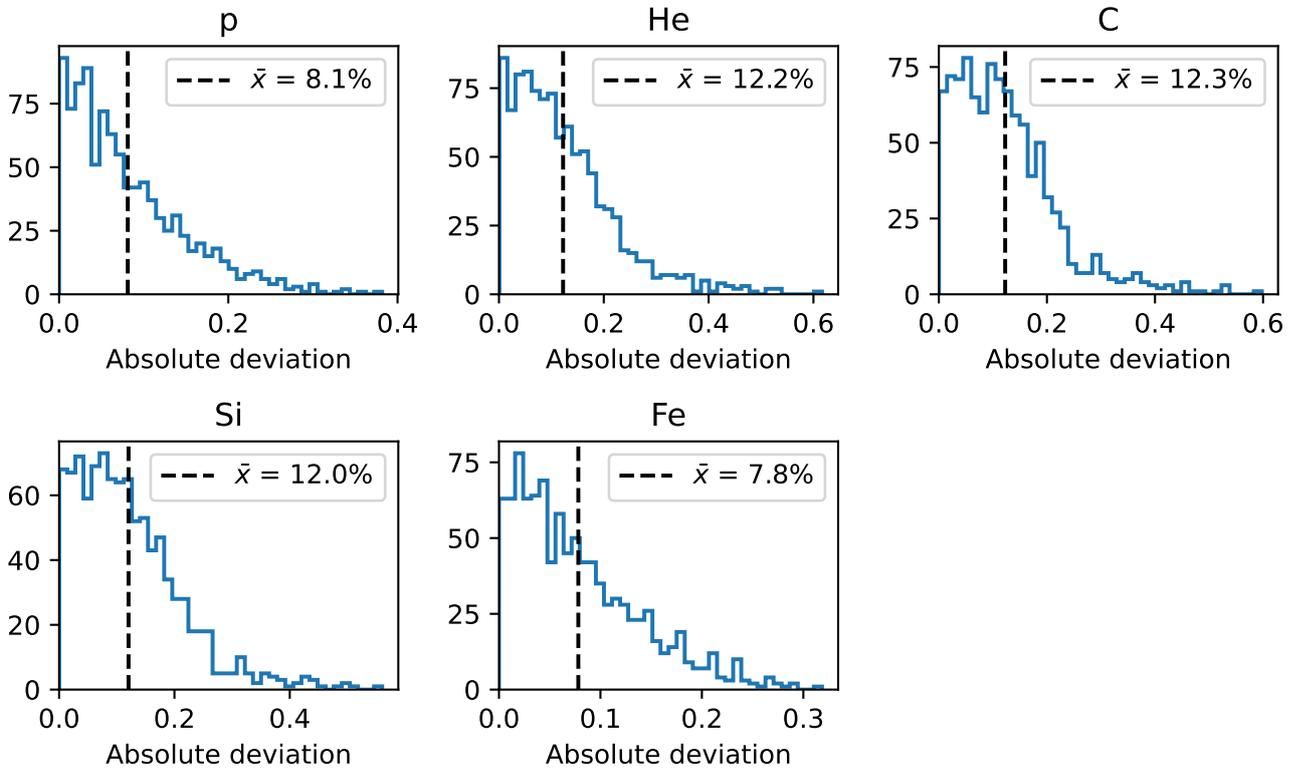


Рисунок 2.5 — Распределение числа событий по абсолютным отклонениям предсказанной CNN доли типа первичной частицы от истинной, заложенной в смесь. Вертикальными линиями показаны средние абсолютные отклонения. CNN обучена и протестирована на Монте-Карло с генератором QGSJet-II.04

Из показанных на рисунке распределений извлекаются средние абсолютные отклонения для каждого типа первичной частицы, и сравниваются эти значения между различными классификаторами. Чем ниже значение, тем лучше классификатор в среднем предсказывает долю компоненты в случайной смеси, а это как раз то, что необходимо от классификатора в исследуемой задаче по восстановлению спектров массовых компонент. Средние абсолютные отклонения для трёх обученных классификаторов для каждого типа первичной частицы приведены в Таблице 1.

Как видно из таблицы, результаты классификаторов близки, однако CNN показывает себя немного лучше остальных моделей. Таким образом, с учётом проведённой проверки на случайных смесях, а также руководствуясь матрицами

Таблица 1 — Средние абсолютные отклонения между предсказанными с помощью классификаторов и заложенными в смесях долями компонент первичных частиц и их статистические погрешности. Все значения приведены в процентах. Все классификаторы обучались и тестировались на Монте-Карло с генератором QGSJet-II.04

Классификатор	p	He	C	Si	Fe
RF	$8,2 \pm 0,2$	$12,5 \pm 0,3$	$12,8 \pm 0,3$	$12,7 \pm 0,3$	$8,6 \pm 0,2$
CNN	$8,1 \pm 0,2$	$12,2 \pm 0,3$	$12,3 \pm 0,3$	$12,0 \pm 0,3$	$7,8 \pm 0,2$
MLP	$9,3 \pm 0,2$	$13,2 \pm 0,3$	$13,5 \pm 0,3$	$13,3 \pm 0,4$	$8,8 \pm 0,2$

смешивания, показанными в Разделе 2.2.1, CNN выбирается основным классификатором, который будет использован для дальнейших исследований.

2.3 Исследование классификатора типа первичной частицы

В этом разделе детально исследуется, как работает CNN классификатор первичной частицы, который был выбран как основной (среди RF, MLP и CNN) ранее. Здесь пойдёт речь о том, какое влияние на CNN оказывают входные параметры; как зависят предсказания классификатора от реконструированной энергии и зенитного угла первичной частицы; насколько изменится качество результатов, если обучать CNN на Монте-Карло с одним генератором, а тестировать с другим; и прочие вопросы, связанные с попытками установить, насколько физически корректно работает CNN.

2.3.1 Исследование важности отдельных параметров

Далее в работе проводится исследование поведения CNN классификатора в зависимости от набора подаваемых на вход параметров (в англоязычной литературе именуемое как *ablation study*). Это необходимо, для того чтобы понимать,

на какие признаки реагирует классификатор, откуда извлекает дискриминирующую информацию и насколько эти признаки физичны.

Перед тем как приступить к изучению, необходимо напомнить, какие входные параметры используются в данной модели машинного обучения и пояснить причины их первоначального выбора. CNN принимает на вход e/γ и μ энерговыделения с детекторных станций, а также набор реконструированных параметров $\log_{10} N_e$, $\log_{10} N_\mu$, θ , s для каждого события.

Для разъяснения выбранных параметров, рассмотрим модель развития ШАЛа Гайтлера-Мэттьюза [78], которая несмотря на свою простоту, верно демонстрирует основные физические характеристики ливня. На примере протона, развитие ШАЛа происходит следующим образом: первичный протон взаимодействует с веществом атмосферы, в результате этого, главным образом, рождаются основной ядерный фрагмент, а также множество заряженных и нейтральных пионов (в равных пропорциях $\pi^\pm : \pi^0$ 2 : 1). Нейтральные пионы, распадаясь согласно $\pi^0 \rightarrow 2\gamma$, запускают электромагнитный ливень. Заряженные пионы, в свою очередь, запускают адронный каскад, продолжая взаимодействовать с веществом атмосферы, и в каждом таком взаимодействии порождают $2N_\pi$ заряженных и N_π нейтральных пионов (например, для пионов с энергиями от 1 ГэВ до 10 ТэВ $N_\pi = 5$ [79]). Развитие адронного ливня происходит до тех пор, пока энергия заряженных пионов не уменьшится до критической, при которой вероятность их распада на мюон и нейтрино $\pi^+ \rightarrow \mu^+ + \nu$ станет выше вероятности взаимодействия с веществом.

Первичные частицы с массовым числом A и энергией E в рамках данной модели рассматриваются как суперпозиция A нуклонов с энергией E/A каждого. Это приводит к тому, что для более тяжёлых первичных частиц (при прочих равных) будет характерна меньшая глубина максимума ливня (X_{max}). А также, из-за более быстрого развития ливней массивных первичных частиц, пионы в адронном каскаде достигнут своей критической энергии (при которой они распадаются, а не взаимодействуют) раньше и, следовательно, увеличат относительное число мюонов по отношению к электромагнитной компоненте.

Таким образом, использование $\log_{10} N_e$, $\log_{10} N_\mu$, а также e/γ и μ энерговыделения с детекторных станций для дискриминации типа первичной частицы по соотношению мюонной и электромагнитной компонент вытекает из представленной выше теории развития ШАЛа. Параметр s , который обозначает «возраст» зарегистрированного ливня, т. е. степень его развития, используется, поскольку

он связан с описанным выше параметром X_{max} . А X_{max} в свою очередь зависит от массы первичной частицы. Наконец, использование θ указывает на толщину атмосферы, в которой распространялся ШАЛ, что также влияет на характеристики ливня.

Дополнительно нужно отдельно отметить, что в качестве входного параметра в CNN не используется матрица времён прихода, поскольку при сравнении экспериментальных данных с Монте-Карло было обнаружено, что калибровка абсолютных значений данного параметра не совпадает.

Для исследования было обучено несколько версий CNN с различным набором входных параметров, например, версия без реконструированных признаков ШАЛ, версия без использования энерговыделений со станций и так далее, и сравнивались результаты каждой из них. Матрицы смешивания для следующих вариантов CNN: стандартной версии; версии, обученной только на e/γ , μ энерговыделениях со станций; версии, обученной только на реконструированных параметрах $\log_{10} N_e$, $\log_{10} N_\mu$, s , θ показаны на Рисунке 2.6.

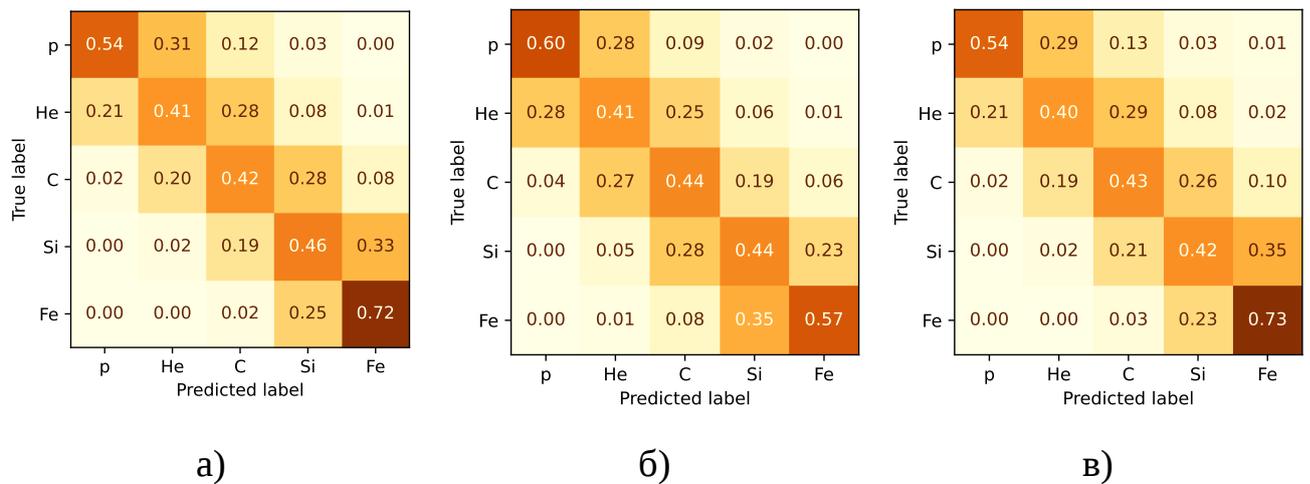


Рисунок 2.6 — Матрицы смешивания для CNN (а) стандартной версии, (б) обученной только e/γ , μ энерговыделениях со станций, (в) обученной только на реконструированных параметрах $\log_{10} N_e$, $\log_{10} N_\mu$, s , θ . В обучении и тесте всех представленных моделей использовано моделирование Монте-Карло с генератором QGSJet-II.04

Видно, что результаты версии нейронной сети, обученной на реконструированных параметрах, практически не отличаются от стандартной версии, однако их матрица смешивания отличается от той, что получена для версии, использующей только энерговыделения со станций. Было выяснено, что основную роль

в этом отличии играет использование реконструированного зенитного угла θ , поскольку он извлекается из измерений времён прибытия со станций KASCADE Array (которые не используются в качестве входных параметров сети) и оказывает влияние на паттерн зарегистрированного ШАЛа.

Таким образом, была обучена CNN с входными параметрами e/γ , μ энерговыделения со станций и реконструированным параметром θ , чьи результаты демонстрируются на Рисунке 2.7. Матрица смешивания этой версии CNN сравнима со стандартной версией и версией на реконструированных параметрах (показанных на Рисунке 2.6). Следовательно, θ является уникальным параметром, который необходимо использовать в классификаторе для максимального извлечения информации о первичной частице из зарегистрированного ШАЛа. Подробное исследование качества классификации в зависимости от зенитного угла θ будет описано ниже в разделе 2.3.2.

Также проверялись результаты классификации такими вариантами CNN, которые не используют данные о числе электронов и e/γ компоненту энерговыделений со станций, и о числе мюонов и μ компоненту. Матрицы смешивания таких классификаторов также показаны на Рисунке 2.7.

Как видно, результаты таких версий CNN существенно деградировали, и поскольку с физической точки зрения без знания и об электромагнитной

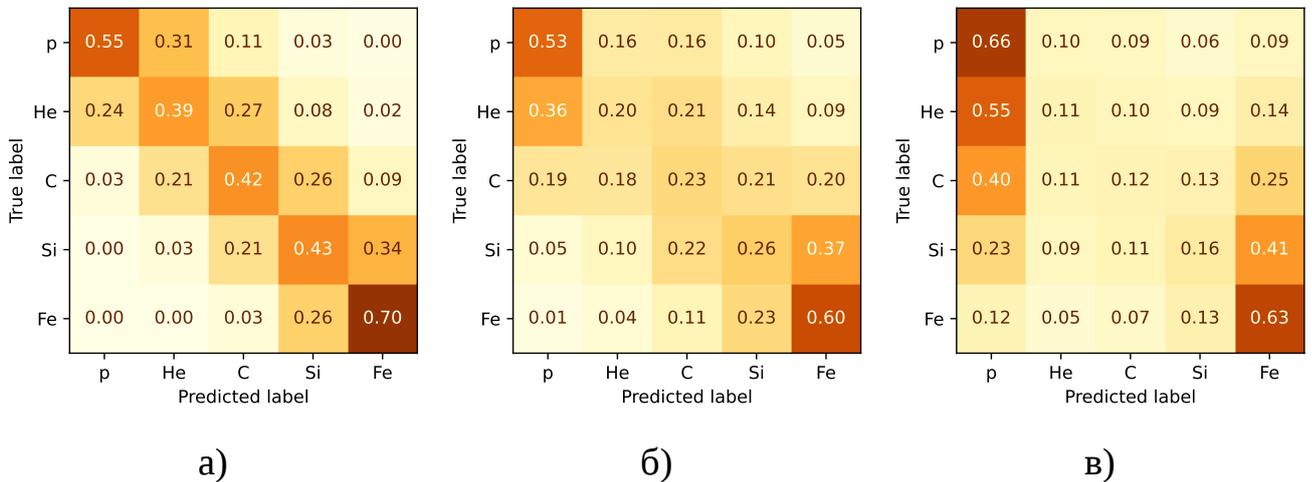


Рисунок 2.7 — Матрицы смешивания для CNN (а) обученной на e/γ , μ энерговыделениях со станций и реконструированном зенитном угле θ , (б) обученной на e/γ энерговыделениях со станций и реконструированных параметрах $\log_{10} N_e$, s , θ , (в) обученной на μ энерговыделениях со станций и реконструированных параметрах $\log_{10} N_\mu$, s , θ В обучении и тесте всех представленных моделей использовано моделирование Монте-Карло с генератором QGSJet-II.04

и о мюонной компоненте ШАЛа разделение по типу первичной частицы сложно, то классификатор ведёт себя в данном случае физически корректно.

В результате исследования степени влияния отдельных входных параметров классификатора выяснилось, что особую важность использования в CNN представляет реконструированный зенитный угол θ , поскольку его невозможно извлечь из других подаваемых параметров, однако он влияет на сигнатуру зарегистрированного ШАЛа. Также была отмечена взаимозаменяемость параметров e/γ энерговыделений со станций и реконструированного числа электронов N_e в ШАЛе, а также μ энерговыделений и N_μ . Однако использование всех этих параметров вместе обеспечивает ускорение сходимости во время обучения данного CNN классификатора.

2.3.2 Зависимость от реконструированной энергии и зенитного угла первичной частицы

Важно рассмотреть, как зависят предсказания CNN классификатора типа первичной частицы от различных реконструированных параметров, например от реконструированной энергии первичной частицы E или от зенитного угла θ . Исследование такой зависимости — это дополнительная проверка качества классификатора CNN, и какое-либо необъяснимое физически поведение классификатора может быть истолковано как возможная ошибка в процессе его разработки и обучения.

Из продемонстрированных на Рисунке 2.8 матриц смешивания

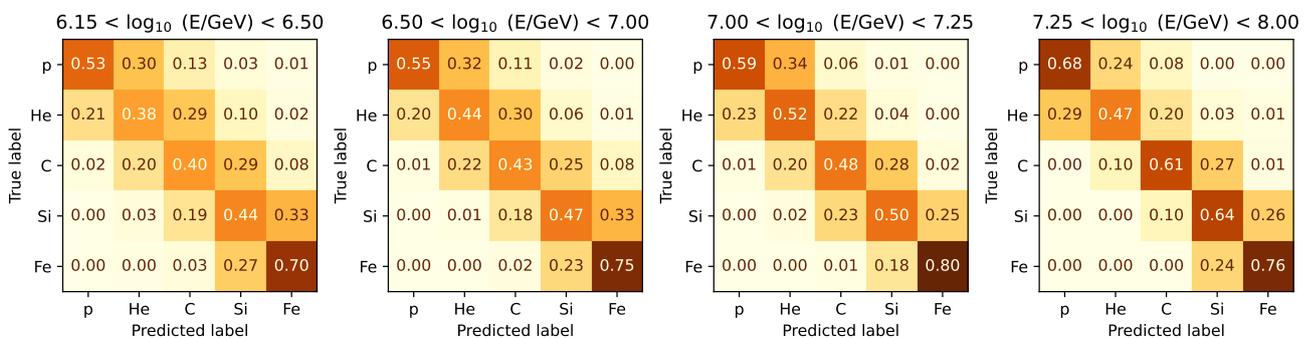


Рисунок 2.8 — Матрицы смешивания CNN для различных диапазонов по реконструированной энергии. Нейронная сеть обучалась и тестировалась на Монте-Карло с генератором QGSJet-II.04

для различных диапазонов по реконструированной энергии первичной частицы E для стандартной CNN видно, что качество классификации в целом растёт с увеличением E , а также, что нет каких-либо выбросов в представленных матрицах. Улучшение классификации с ростом энергии можно объяснить тем, что при более высоких энергиях первичной частицы, высота, на которой достигается максимальное развитие ШАЛа, становится ниже и ШАЛы покрывают большую площадь, тем самым, энерговыделения регистрируются в большем количестве станций, как следствие, сигнатура таких ливней чётче и возможность их покомпонентного разделения выше.

Чтобы более аккуратно продемонстрировать зависимость качества классификации от энергии, были построены зависимости диагональных элементов матрицы смешивания от реконструированной энергии, которые показаны на Рисунке 2.9, поскольку в данном случае диагональные элементы в наибольшей степени показывают степень разделения массовых компонент.

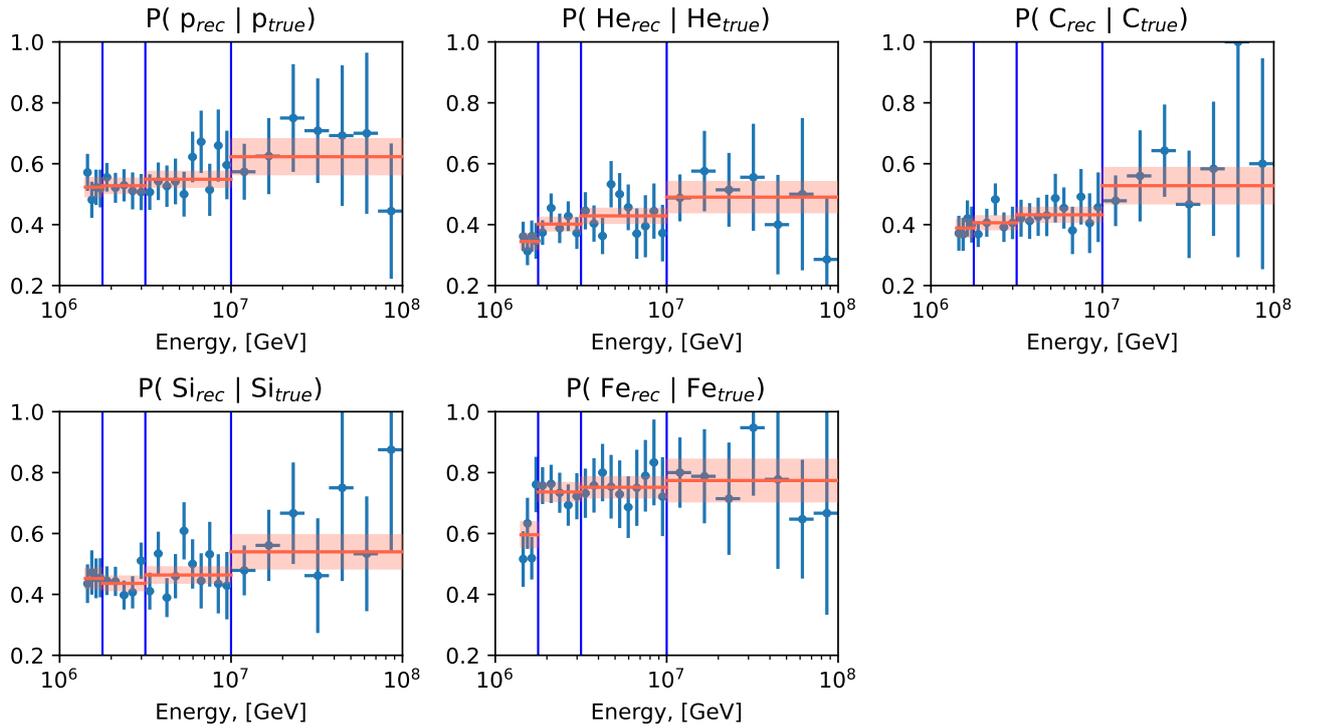


Рисунок 2.9 — Зависимость диагональных элементов матриц смешивания от реконструированной энергии E для CNN, обученной и тестированной с генератором адронных взаимодействий QGSJet-II.04. Синие вертикальные линии обозначают диапазоны, в которых Монте-Карло объединяется для последующей процедуры анфолдинга. Красные горизонтальные линии показывают средние значения данных матричных элементов в ограниченных диапазонах энергий вместе с их неопределённостями

Можно отметить, что наблюдается улучшение качества дискриминации с увеличением E , однако также видно, что статистические ошибки с ростом энергии увеличиваются, что объясняется недостаточным количеством событий Монте-Карло моделирования, особенно при высоких энергиях (т.к. в Монте-Карло заложен спектр $dN/dE \propto E^{-2}$), поэтому для последующей процедуры анфолдинга было решено объединить данные в несколько энергетических диапазонов по $\log_{10}(E/\text{ГэВ})$: 6,15–6,25, 6,25–6,5, 6,5–7, 7–8. В этих диапазонах матрицы ведут себя стабильно и, таким образом, объединение не окажет влияние на анфолдинг, что также показано на Рисунке 2.9.

Хорошая проверка для классификатора — это представленное на Рисунке 2.10 исследование поведения предсказанного им массового состава от реконструированного зенитного угла θ , поскольку известно, что в идеальном случае

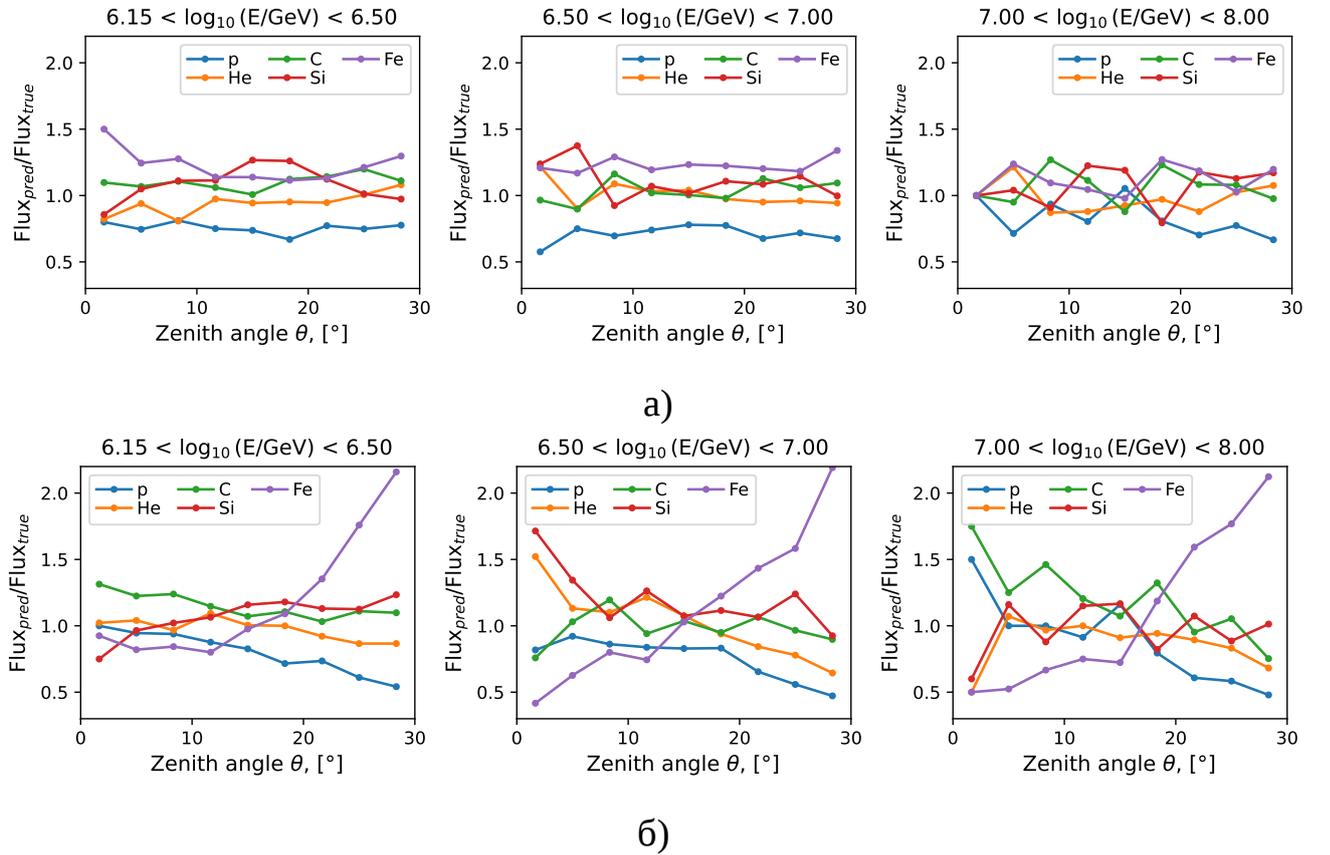


Рисунок 2.10 — Отношение предсказанного потока ($\text{Flux}_{\text{pred}}$) к истинному ($\text{Flux}_{\text{true}}$) в зависимости от реконструированного зенитного угла θ для пяти массовых компонент в трёх энергетических диапазонах. На рисунках показаны результаты для (а) стандартной CNN, принимающей на вход e/γ , μ энергосодержания, θ , $\log_{10} N_e$, $\log_{10} N_\mu$, s , и (б) версии CNN без θ входного параметра (остальные как в стандартной CNN). Обучение и проверка проводились на Монте-Карло с генератором QGSJet-II.04

доля событий каждой массовой компоненты не должна зависеть от θ . В продолжение исследования влияния отдельных параметров на классификатор из раздела 2.3.1 было проведено сравнение отношения предсказанного потока числа частиц к истинному, или, в данном случае эквивалентно, отношения предсказанных долей массовых компонент к реальным (заложенным в Монте-Карло) в зависимости от θ для стандартной CNN, использующей в т.ч. θ в качестве входного параметра, и для версии CNN, не использующей θ .

Из рисунка 2.10 видно, что стандартный классификатор, использующий θ , не содержит каких-либо явных зависимостей, в то время как CNN без использования θ демонстрирует значительный прирост доли предсказанной железной компоненты с увеличением зенитного угла и, таким образом, использование угла θ в качестве входного параметра нейронной сети оправдано.

Также исследовалось поведение стандартного CNN классификатора в двух различных диапазонах зенитных углов — $\theta < 18^\circ$ и $\theta < 30^\circ$. Матрицы смешивания в этих случаях показаны на Рисунке 2.11, а отношения предсказанного потока к реальному как функция от реконструированной энергии первичной частицы продемонстрированы на Рисунке 2.12.

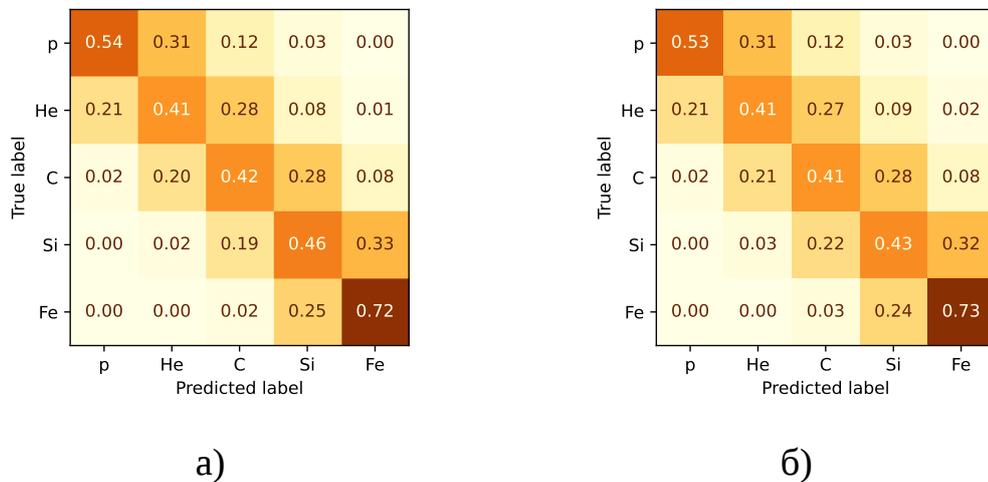


Рисунок 2.11 — Матрицы смешивания CNN для отборов по зенитному углу (а) $\theta < 18^\circ$ и (б) $\theta < 30^\circ$. CNN обучалась и тестировалась на Монте-Карло с генератором адронных взаимодействий QGSJet-II.04

Заметного ухудшения результатов для более слабого ограничения на θ не наблюдается, поэтому для задач, в которых необходима дополнительная статистика экспериментальных данных, ослабленный отбор на зенитный угол принесёт пользу. Однако в данном исследовании по изучению массового состава расширение отбора не приведёт к какому-либо улучшению, поскольку, как будет

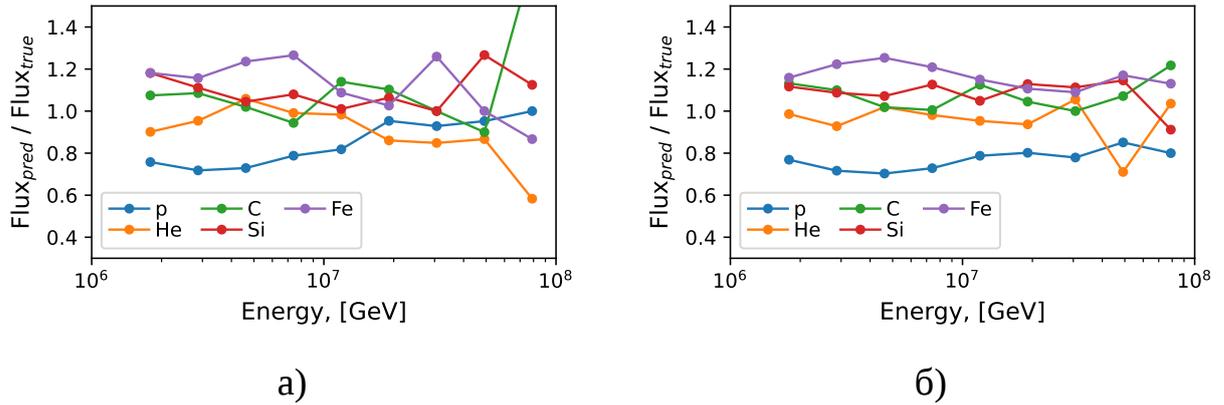


Рисунок 2.12 — Зависимость отношения предсказанного CNN потока к реальному от реконструированной энергии для пяти массовых компонент для следующих критериев отбора по зенитному углу: (а) $\theta < 18^\circ$, (б) $\theta < 30^\circ$. CNN обучена и тестирована на Монте-Карло с генератором адронных взаимодействий QGSJet-II.04

обсуждаться ниже, над статистическими погрешностями доминируют систематические погрешности.

2.3.3 Неработающие детекторные станции

В Монте-Карло моделировании в каждом событии все детекторные станции KASCADE Array стабильно работают, однако в реальных экспериментальных данных это не так. Примеры усреднённых e/γ энерговыделений со станций для нескольких заходов в экспериментальных данных показаны на Рисунке 2.13. Из этого рисунка видно, что в части экспериментальных заходов есть неработающие детекторные станции, которые показаны тёмными областями с нулевым энерговыделением. Таким образом, поскольку в данных присутствуют ШАЛы, которые регистрировались, когда некоторые станции не работали (в отличие от Монте-Карло), то необходимо учесть влияние таких событий на CNN классификатор типа первичной частицы.

Во-первых, необходимо оценить число событий с неработающими станциями, а также количество таких станций в событии. Для этого e/γ энерговыделения со станций в эксперименте усреднялись в пределах одного захода (run) и средние значения сравнивались с ожидаемыми. Оказалось, что порядка $\sim 50\%$

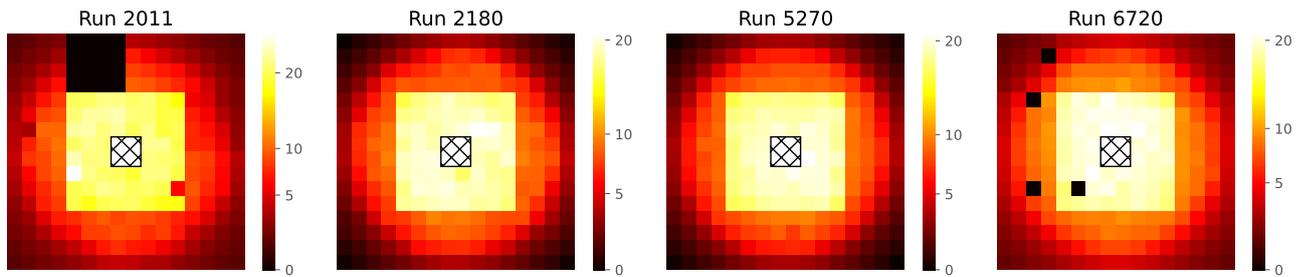


Рисунок 2.13 — Примеры усреднённых значений e/γ энерговыделений в МэВ для экспериментальных заходов (слева направо) 2011, 2180, 5270, 6720. В KASCADE Array нет детекторных станций в центральной области 2×2 , эта область показана штриховкой. Усреднённые значения энерговыделений во внутренней области заметно ярче, поскольку в станциях типа II (расположенных во внутренней 8×8 области) установлено по 4 e/γ детектора, в то время как в станциях типа I только по 2

заходов содержат неработающие станции, что достаточно много и должно, таким образом, учитываться в анализе.

Далее, нужно учесть влияние таких событий, и общая идея состоит в следующем — «испортим» события тестовой части Монте-Карло и пропустим их через используемый CNN классификатор, сравним полученные результаты со стандартными и в дальнейшем пересчитаем эту разницу в соответствующую систематическую погрешность (подробнее в разделе 2.5). Однако «испортить» возможно лишь энерговыделения с детекторных станций, просто занулив часть из них, но не реконструированные параметры (N_e , N_μ , θ , s), которые также использует CNN, поскольку реконструкция этих параметров в Монте-Карло, проводится на стороне KCDC. Поэтому в данной задаче по определению степени влияния неработающих станций, использовалась CNN, обученная только на e/γ и μ энерговыделениях. И хотя матрица смешивания такой версии, как было выяснено в разделе 2.3.1, отличается от стандартной, но её деградация на «испорченном» Монте-Карло всё-таки будет указывать на эффект неработающих станций.

Таким образом, для классификации использовались «испорченные» Монте-Карло события, в которых занулялось e/γ и μ энерговыделение в случайной области из 4×4 станций. Пример такого события показан на Рисунке 2.14. На Рисунке 2.15 показано сравнение матриц смешивания на обычном и на «испорченном» Монте-Карло моделировании.

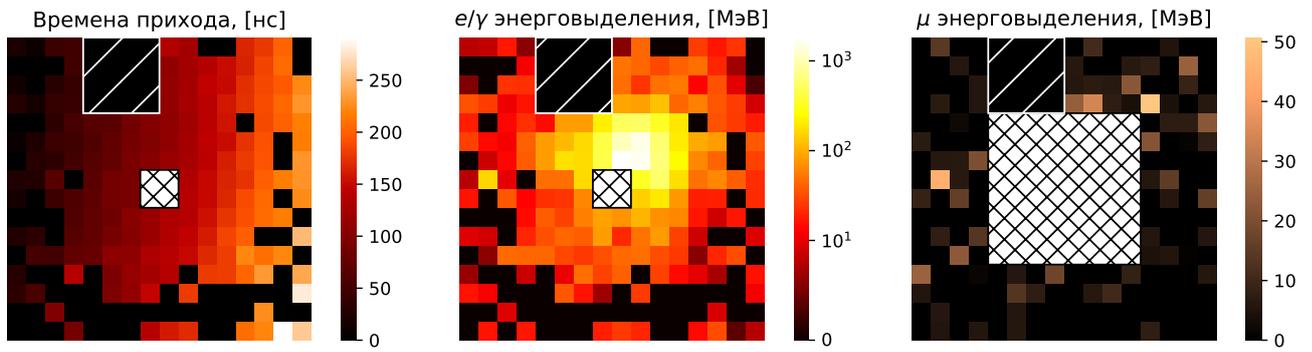
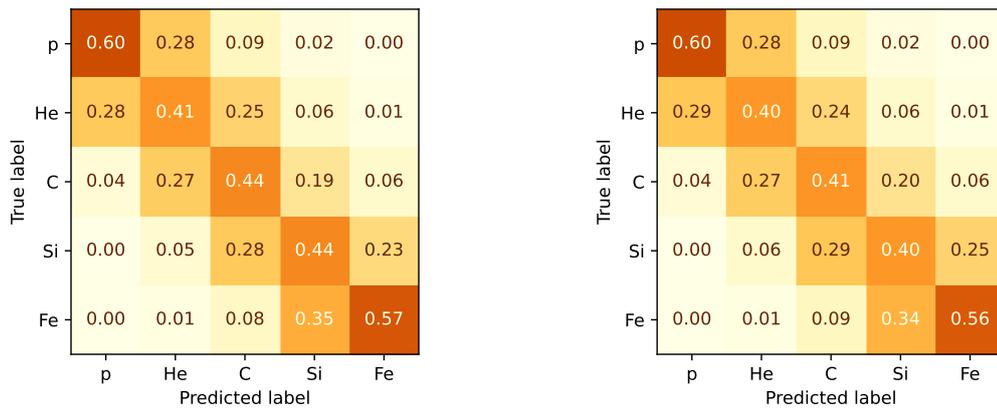


Рисунок 2.14 — Пример «испорченного» события из Монте-Карло. Область за- нуления энергосвечений показана белой штрихованной областью



а)

б)

Рисунок 2.15 — Матрицы смешивания для CNN, использующих в качестве вход- ных параметров только e/γ , μ энергосвечения: (а) с использованием для обуче- ния и теста стандартного Монте-Карло моделирования с генератором адронных взаимодействий QGSJet-II.04, (б) то же, но для теста использовано «испорчен- ное» Монте-Карло моделирование

Как видно из рисунка 2.15, средняя деградация диагональных элементов составляет $\sim 2\%$, что выглядит не критично, однако далее необходимо пересчи- тать и учесть это ухудшение в соответствующей систематической неопределён- ности, этот процесс будет подробно описан в разделе 2.5.2.

2.3.4 Зависимость от генераторов адронных взаимодействий

Как было отмечено, изучение космических лучей с энергиями выше 10^{14} эВ происходит посредством регистрации ШАЛов. Физические выводы в таких

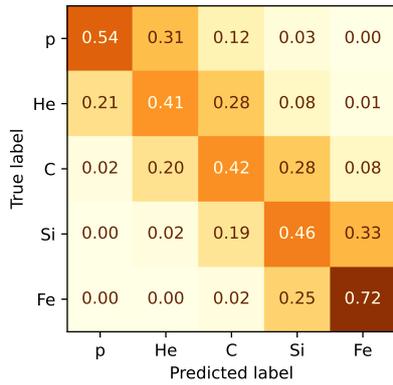
экспериментах можно делать только при сопоставлении распределений наблюдаемых величин с предсказанными на основе Монте-Карло моделирования, причём именно степень надёжности Монте-Карло на сегодняшний день служит источником наибольшей неопределенности. Более конкретно, к ней приводят различия в предсказаниях наблюдаемых для различных генераторов адронных взаимодействий высоких энергий, которые отвечают за взаимодействие первичной частицы с атмосферой и начальное развитие ШАЛа.

В настоящее время активно развивается несколько независимых генераторов адронных взаимодействий высоких энергий: QGSJet, EPOS, Sibyll, которые были подстроены под различные результаты ускорительных экспериментов и выдержали различные тесты в экспериментах с ШАЛами.

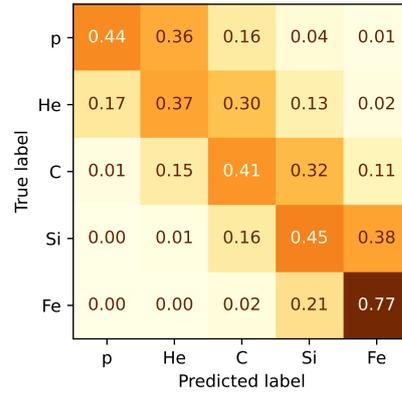
Современные (post-LHC) версии этих генераторов (по сравнению с предыдущими pre-LHC версиями) используют информацию экспериментов на Большом адронном коллайдере, в частности информацию о сечении упругого рассеяния pp и полном сечении pp , полученную в экспериментах TOTEM [80] и ATLAS [81], что позволяет ограничить предсказания генераторов относительно начальной точки развития атмосферного ливня [82]. Однако хотя различия между генераторами были уменьшены, но они всё ещё представляют существенную проблему, открытые вопросы касаются учёта пионных взаимодействий, которые оказывают влияние на геометрию и энергию мюонов в ШАЛах, что может быть учтено в следующих версиях генераторов [37].

Поскольку в нашем арсенале есть наборы данных Монте-Карло, основанных на трёх независимо развивающихся генераторах адронных взаимодействий высоких энергий, каждый из которых претендует на истинность, то интересно сравнить, как они соотносятся друг с другом в данной задаче определения массового состава. Так, проведено сравнение матриц смешивания для классификаторов типа первичной частицы, обученных или протестированных с помощью Монте-Карло моделирования, основанного на каждом из трёх различных post-LHC версиях генераторов из этих семейств, а именно QGSJet-II.04, EPOS-LHC, Sibyll 2.3c.

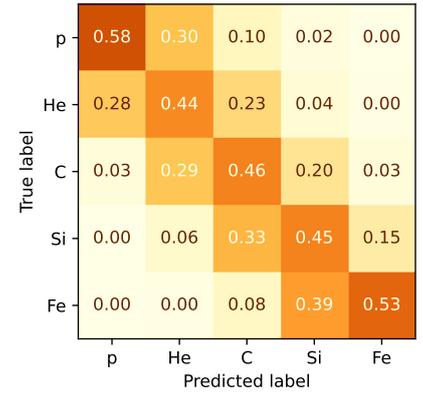
В результате в дополнение к стандартной CNN (использующей Монте-Карло с QGSJet-II.04 в качестве тренировочного) были использованы версии той же CNN на Монте-Карло с EPOS-LHC и Sibyll 2.3c. Далее все версии CNN тестировались на тестовых наборах Монте-Карло всех версиях генераторов. Матрицы смешивания с результатами представлены на Рисунке 2.16.



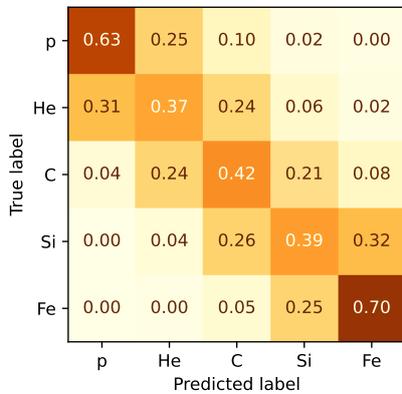
а) Q, Q



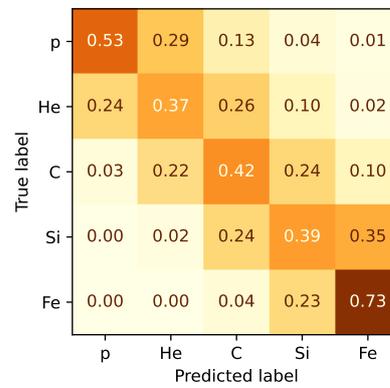
б) Q, E



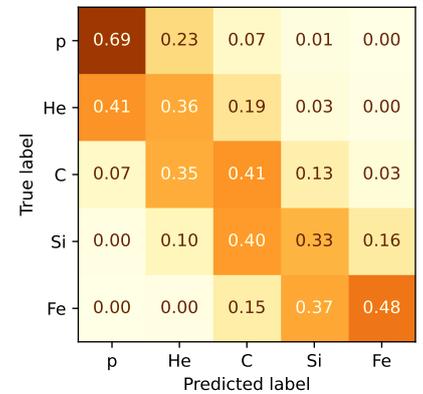
в) Q, S



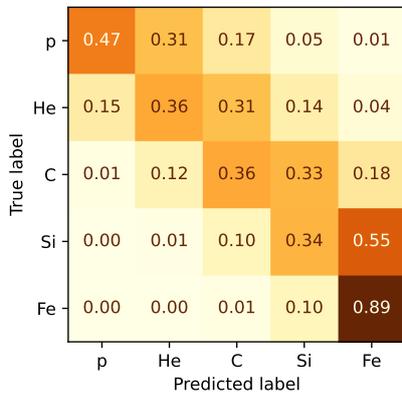
г) E, Q



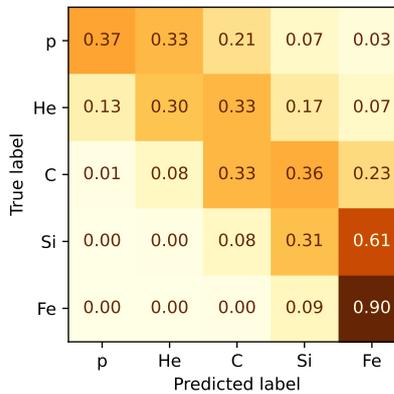
д) E, E



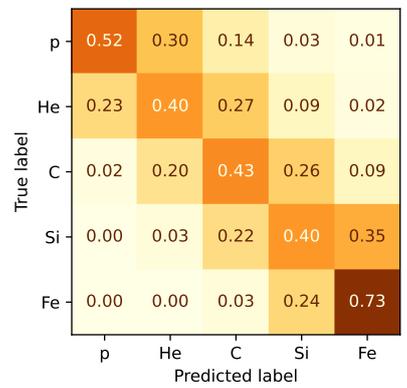
е) E, S



ж) S, Q



з) S, E



и) S, S

Рисунок 2.16 — «Кросс-адронные» матрицы смешивания для CNN классификатора, обученного и протестированного на Монте-Карло с тремя различными генераторами адронных взаимодействий QGSJet-II.04, EPOS-LHC, Sibyll 2.3с, которые обозначены Q, E, S соответственно для краткости. В подписях к матрицам сначала указан генератор, используемый в обучении, после запятой — на тесте

Как уже было отмечено ранее, качество классификации, полученное на тех же генераторах, на которых и были обучены версии CNN, в целом похоже; наибольших значений у этих матриц достигают диагональные элементы, нет каких-либо артефактов; максимальное смешивание происходит с соседними массовыми компонентами. Также видно, что так называемые «кросс-адронные» матрицы (полученные при обучении CNN на Монте-Карло с одним генератором, а тесте — с другим), имеют некоторый сдвиг, так например Sibyll 2.3с выглядит более «железной» моделью, поскольку классификатор на его основе систематически пытается определить компоненты в других моделях более тяжёлыми, чем они есть на самом деле; аналогично, EPOS-LHC среди трёх исследуемых моделей выглядит более «протонной».

Таким образом, поскольку уже на данном этапе видно, что «кросс-адронные» матрицы смешивания заметно отличаются для различных семейств исследованных современных (post-LHC) адронных генераторов и говорит о сохраняющемся несоответствии между ними, то при определении энергетических спектров индивидуальных массовых компонент необходимо учесть вклад от их несоответствия, подробно это будет рассмотрено в разделе 3.2.

Однако как уже утверждалось, реальные ШАЛы неточно описываются всеми генераторами, с этим связана, например, «мюонная загадка» [83] — значительный избыток числа мюонов в реальных ливнях относительно предсказываемого в Монте-Карло. И, следовательно, разница между моделями даёт лишь условную оценку реальной неопределенности.

2.3.5 Спектры индивидуальных массовых компонент

Т. к. изучаемый CNN классификатор событийно реконструирует тип первичной частицы, то построим спектры реконструированных массовых компонент ($\frac{dJ}{dE}$) в зависимости от реконструированной энергии первичной частицы (E) согласно формуле

$$\frac{dJ}{dE}(E_i) = \frac{N_i}{\mathcal{E} \Delta E_i},$$

где i — порядковый номер бина по реконструированной энергии, N_i — количество событий в i -м энергетическом бине, \mathcal{E} — экспозиция (подробности

в разделе 1.2), ΔE_i — ширина бина. Такие спектры будут называться «свёрнутыми», т. е. предсказанные напрямую классификатором без поправок на матрицу смешивания. Полученные на открытой части экспериментальных данных «свёрнутые» спектры индивидуальных массовых компонент реконструированные классификатором CNN, обученным с генератором QGSJet-II.04, показаны на Рисунке 2.17.

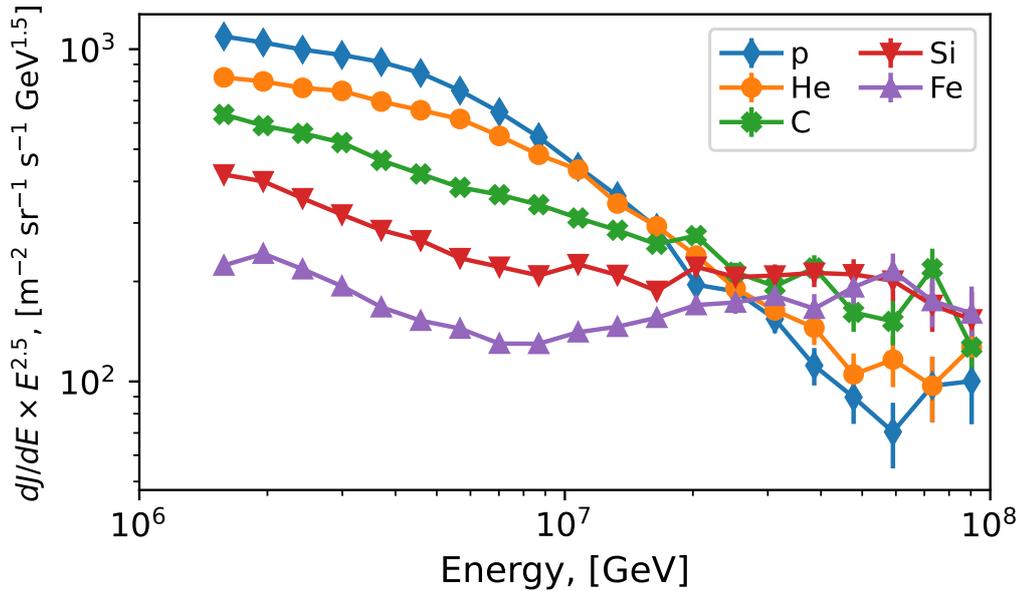


Рисунок 2.17 — «Свёрнутые» спектры пяти индивидуальных массовых компонент (p , He , C , Si , Fe) в зависимости от реконструированной энергии. CNN классификатор обучен на Монте-Карло с генератором QGSJet-II.04. Вертикальными линиями показаны статистические погрешности

Как было указано ранее, исследование проводится в диапазоне энергий $10^{6,15} - 10^8$ ГэВ, а по аналогии с KASCADE полный интервал был разбит на 20 эквидистантных бинов по логарифму энергии. Также здесь и далее спектры демонстрируются с дополнительным множителем $E^{2.5}$ (где E — энергия, отложенная по оси абсцисс) для лучшего визуального восприятия их формы и поведения.

Из рисунка можно заметить, что спектры достаточно гладкие в зависимости от энергии, хотя CNN не получает напрямую информацию об этой переменной; также отсутствуют какие-либо артефакты и выбросы, что добавляет уверенности в качестве реконструкции. Также видно качественное поведение — если на более низких энергиях доли массовых компонент уменьшаются в порядке увеличения их массовых чисел, то на энергиях выше ~ 20 ПэВ происходит кардинальное изменение порядка компонент. Однако каких-либо более конкретных фактов ещё нельзя извлечь, поскольку при текущей точности классификатора

CNN (около 50 %), а также разрешения стандартной реконструкции энергии первичной частицы ($\sim 10\%$) требуется анфолдинг спектров, о котором пойдёт речь в следующей части.

2.4 Анфолдинг

Процесс анфолдинга — это процедура восстановления оригинального распределения по наблюдаемому, «загрязнённому» какими-либо эффектами. В случае данной работы, на текущей стадии имеются «свёрнутые» спектры, т.е. распределение числа событий по реконструированным массовым компонентам и по реконструированной энергии первичной частицы. Однако поскольку, как было отмечено выше, ни реконструкция энергии, ни реконструкция типа первичной частицы не выполнена с идеальной точностью, то необходимо провести процедуру анфолдинга для восстановления распределения событий по реальной массовой компоненте и энергии первичной частицы.

С математической точки зрения, проблема анфолдинга в случае дискретного распределения состоит в определении вектора ожидаемого числа событий в оригинальном распределении (μ_i) по вектору ожидаемого числа событий в наблюдаемом распределении (ν_j) и известной матрице отклика R_{ij} , которые связаны следующим образом,

$$\nu_i = \sum_{j=1}^M R_{ij} \mu_j, i = 1, \dots, N, \quad (2.1)$$

где N/M — число бинов в наблюдаемом / оригинальном распределении. Кроме того нужно отметить, что элементы матрицы отклика в терминах условных вероятностей записываются так, $R_{ij} = P(\text{obs. } j | \text{true } i)$, т.е. как вероятность наблюдать событие в бине i при условии, что в оригинальном распределении событие попало в бин j . Таким образом, матрица отклика R повторяет матрицу смешивания классификатора в случае анфолдинга по типу массовой компоненты с точностью до транспонирования.

В данном случае необходимо провести анфолдинг по двум переменным: по реконструированной массовой компоненте ($A_m^{\mathcal{R}}$) и по реконструированной

энергии первичной частицы ($E_n^{\mathcal{R}}$), чтобы определить $\mu(A_i^{\mathcal{T}}, E_j^{\mathcal{T}})$ — вектор с распределением числа событий по оригинальным массовым компонентам $A_i^{\mathcal{T}}$ и по оригинальной энергии первичной частицы $E_j^{\mathcal{T}}$.

Сначала рассмотрим анфолдинг по массовой компоненте. Согласно формуле полной вероятности, вероятность реконструировать массовую компоненту события в бине по реконструированной энергии записывается следующим образом

$$P(A_i^{\mathcal{R}}|E_m^{\mathcal{R}}) = \sum_j P(A_i^{\mathcal{R}}|A_j^{\mathcal{T}}, E_m^{\mathcal{R}}) \cdot P(A_j^{\mathcal{T}}|E_m^{\mathcal{R}}), \quad (2.2)$$

где $A_i^{\mathcal{R}}$ — i -ая реконструированная массовая компонента первичной частицы, $E_m^{\mathcal{R}}$ — m -ый бин реконструированной энергии первичной частицы. Как видно, это представление может быть переписано в виде Формулы (2.1), и является стандартной задачей анфолдинга, поскольку при известных из Монте-Карло $P(A_i^{\mathcal{R}}|A_j^{\mathcal{T}}, E_m^{\mathcal{R}})$, и полученных из экспериментальных данных $P(A_i^{\mathcal{R}}|E_m^{\mathcal{R}})$ требуется определить $P(A_j^{\mathcal{T}}|E_m^{\mathcal{R}})$.

По аналогии можно записать энергетический анфолдинг в виде

$$P(E_m^{\mathcal{R}}) = \sum_j P(E_m^{\mathcal{R}}|E_j^{\mathcal{T}}) \cdot P(E_j^{\mathcal{T}}), \quad (2.3)$$

в результате которого будет определён вектор $P(E_m^{\mathcal{T}})$ вероятностей обнаружить событие в интервале истинных энергий $E_m^{\mathcal{T}}$. Как показано на Рисунке 1.4, энергия первичной частицы реконструирована с $\sim 10\%$ точностью, поэтому вообще проводится энергетический анфолдинг. А поскольку целью является определить истинное число событий в конкретно заданных энергетических бинах, то реконструированную энергию можно рассматривать в дискретном смысле (попадание энергии события в какой-либо энергетический бин), и, таким образом, построить энергетическую матрицу смешивания (и, следовательно, матрицу отклика для энергий). Для демонстрации, несколько строк из полученной из Монте-Карло моделирования с генератором QGSJet-II.02 энергетической матрицы смешивания представлено на Рисунке 2.18.

Таким образом, скомбинировав два результата, получаем следующее

$$\begin{aligned} \mu(A_i^{\mathcal{T}}, E_j^{\mathcal{T}}) &= \mu_{tot} P(A_i^{\mathcal{T}}, E_j^{\mathcal{T}}) = \\ &= \mu_{tot} P(A_i^{\mathcal{T}}|E_j^{\mathcal{T}}) P(E_j^{\mathcal{T}}) \rightarrow \mu_{tot} P(A_i^{\mathcal{T}}|E_j^{\mathcal{R}}) P(E_j^{\mathcal{T}}), \end{aligned} \quad (2.4)$$

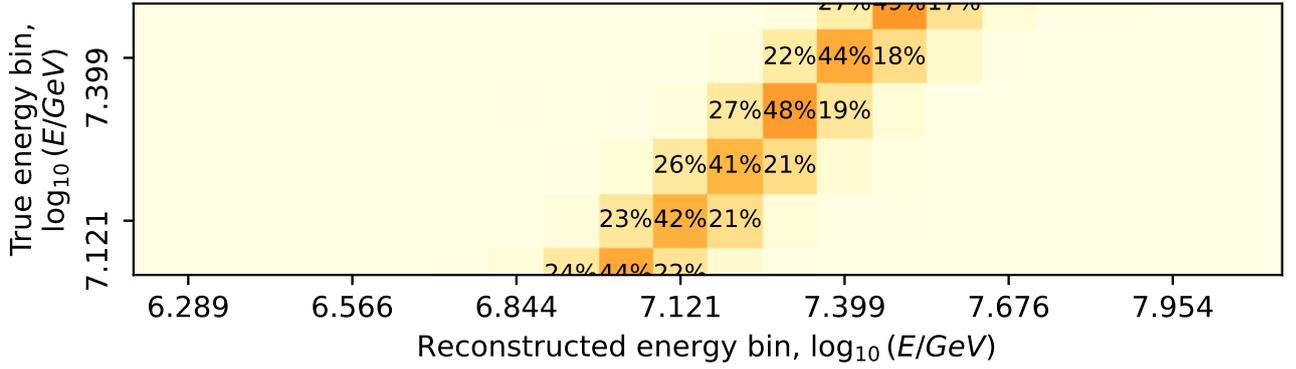


Рисунок 2.18 — Часть энергетической матрицы смешивания для Монте-Карло, основанном на QGSJet-II.02. Численно приведены значения для элементов матрицы с вероятностями выше 10 %

где μ_{tot} — полное число событий, которое не меняется в процессе анфолдинга, поскольку в данном исследовании матрицы отклика фиксируют только перераспределение компонент и энергий. Так же, как видно из Формулы (2.4), производится замена $P(A_i^T | E_j^T)$ на $P(A_i^T | E_j^R)$, чтобы использовать результат анфолдинга по типу массовой компоненты. Эта замена является источником систематической неопределённости, и её влияние будет подробно оцениваться в Разделе 2.5.

Наконец, необходимо корректно учесть эффективность регистрации событий, которая, отличается для нескольких энергетических бинов от единицы. Это можно наблюдать на Рисунке 2.19, который демонстрирует эффективность критериев отбора для различных массовых компонент в зависимости от энергии первичной частицы. В то же время эффективность срабатывания триггера равна единице для используемых отборов. Учёт эффективности проводится так

$$\mu^0(A_i^T, E_j^T) = \frac{\mu(A_i^T, E_j^T)}{\varepsilon(A_i^T, E_j^T)}, \quad (2.5)$$

где μ^0 — скорректированное число событий, ε — эффективность регистрации событий конкретного массового типа A_i^T в заданном энергетическом интервале E_j^T (извлекается из соответствующего Монте-Карло).

Тогда итоговая формула для определения энергетических спектров индивидуальных массовых компонент $\frac{dJ(A_i^T)}{dE_j^T}$ (далее для простоты будет записываться dJ/dE) выглядит следующим образом:

$$\frac{dJ(A_i^T)}{dE_j^T} = \frac{\mu^0(A_i^T, E_j^T)}{\varepsilon \Delta E_j^T}, \quad (2.6)$$

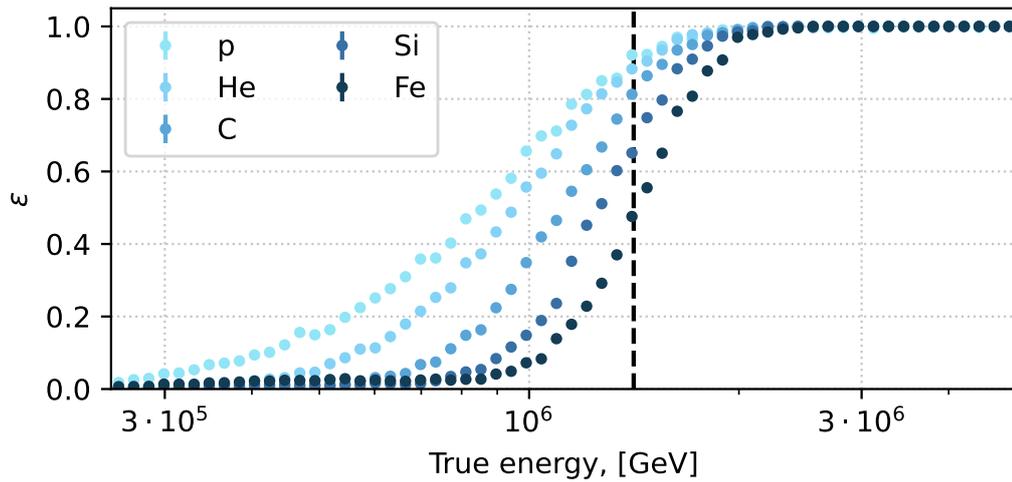


Рисунок 2.19 — Эффективность критериев отбора в зависимости от истинной энергии первичной частицы по массовым компонентам для Монте-Карло с генератором QGSJet-II.04. Вертикальной чёрной линией продемонстрирована минимальная энергия $10^{6,15}$ ГэВ, с которой начинается анализ массового состава в данном исследовании

где \mathcal{E} — экспозиция (подробности в разделе 1.2), ΔE_j^T — ширина энергетического бина.

Для решения описанных выше задач анфолдинга применялся байесовский итеративный метод [84] с использованием python-пакета ruunfold [85]. Анфолдинг проводился по пяти исследуемым массовым компонентам (p , He , C , Si , Fe), и в интервале от $10^{6,15}$ ГэВ до 10^8 ГэВ по энергии на 20 бинах, равномерно расположенных в логарифмическом масштабе. События не попавшие в полный изучаемый энергетический диапазон записывались в дополнительные так называемые нижний и верхний бины, которые учитывались при энергетическом анфолдинге. В качестве критерия остановки использовался χ^2 между векторами значений, полученных на текущей и предыдущей итерации, в случае анфолдинга по типу первичной частицы итерационный процесс останавливался при $\chi^2 = 0,01$, для анфолдинга по энергии — при 0,1.

Статистическая погрешность полученных результатов вычислялась пакетом ruunfold методом распространения ошибки через матрицы отклика стандартной статистической погрешности, связанной с пуассоновской вероятностью обнаружить некоторое число экспериментальных событий в энергетическом и массовом бине.

Также, неопределённости в матрицах отклика, связанные с ограниченным объёмом моделирования Монте-Карло, рассматривались как пуассоновские.

Учёт вклада от критерия остановки и неопределённостей в матрицах отклика в суммарную систематическую погрешность подробно описан в Разделе 2.5.

2.5 Изучение систематических погрешностей

После проведения процедуры анфолдинга необходимым этапом становится оценка систематических погрешностей полученного результата. В данной главе приводится список исследованных систематических погрешностей, а также методы по их оценке. Полную систематическую погрешность для результатов, в основе которых лежит анфолдинг, использующий Монте-Карло моделирование с конкретным генератором адронных взаимодействий, назовём «базовой» для соответствующего генератора. «Базовая» систематическая погрешность вычислялась как квадратичная сумма отдельных погрешностей, описываемых ниже. Учёт погрешностей, связанных с использованием различных наборов Монте-Карло моделирования, основанных на трёх различных post-LHC генераторах (QGSJet-II.04, EPOS-LHC, Sibyll 2.3c), будет описан в Разделе 3.2.

2.5.1 Неопределённости анфолдинга

В эту группу входят неопределённости, связанные с

- регуляризацией процедуры анфолдинга,
- погрешностями значений в матрицах отклика,
- заменой в процедуре анфолдинга (см. Формулу (2.4)).

Регуляризация анфолдинга, которая в данном случае заключается в остановке итерационного процесса согласно критерию χ^2 (как указано в предыдущем разделе).

Использование различных регуляризаций в анфолдинге является стандартной практикой, поскольку позволяет стабилизировать процесс, уменьшая разброс полученных значений, однако взамен приводит к их смещённой оценке. Таким образом, для того чтобы определить степень надёжности полученных результатов, необходимо выяснить величину этого смещения. Поскольку в данной

работе критерий остановки используется как в энергетическом, так и в анфолдинге по типу первичной частицы, то рассматриваются два независимых вклада.

Однако получить точную величину смещения было бы возможно лишь в случае, когда известны истинные значения оцениваемых нами параметров, но тогда анфолдинг был бы не нужен вовсе, поэтому воспользуемся методом приближённой оценки.

С формальной точки зрения задача выглядит так, пусть дан \vec{v} — вектор значений наблюдаемого распределения (по которому необходимо получить оценку истинного), \mathbf{R} — матрица отклика, а также $\vec{\mu}_{est}$ — полученный в результате анфолдинга (вектора \vec{v} матрицей \mathbf{R}) вектор с оценкой значений истинного распределения. Необходимо определить смещение вектора $\vec{\mu}_{est}$ относительно вектора значений истинного распределения $\vec{\mu}$. По определению это смещение \vec{b} записывается как:

$$\vec{b} = \vec{\mu}_{est} - \vec{\mu}$$

Однако значения $\vec{\mu}$ неизвестны, поскольку являются главной целью проведения анфолдинга, поэтому пытаемся определить некоторое приблизительное значение смещения, которое назовём \vec{b}_{est} , и будем считать его систематической неопределённостью из-за регуляризации анфолдинга.

Для этого примем $\vec{\mu}_{est}$ за вектор истинных значений оцениваемых параметров, и проведём над ним процедуру свёртки, т. е. свернём $\vec{\mu}_{est}$ с матрицей отклика \mathbf{R} , определяя таким образом \vec{v}_{est} , и, считая значения этого вектора распределёнными согласно пуассоновскому закону, на основе этого вектора получим N распределений \vec{v}_{est}^i (индекс i от 1 до N). Затем проводим анфолдинг над каждым \vec{v}_{est}^i и в результате, получив векторы $\vec{\mu}_{est1}^i$, вычислим приближённое смещение \vec{b}_{est} , как

$$\vec{b}_{est} = \frac{1}{N} \sum_{i=1}^N (\vec{\mu}_{est1}^i - \vec{\mu}_{est}).$$

Этот метод оценки смещения взят из работы [86], а также похож на метод, используемый в диссертации [62], которая резюмирует оригинальные результаты эксперимента KASCADE по реконструкции массовых компонент.

Проверить метод оценки можно на событиях Монте-Карло, так например на Рисунке 2.20 показано сравнение оценки и реального смещения для спектра протонов и железа в зависимости от энергии.

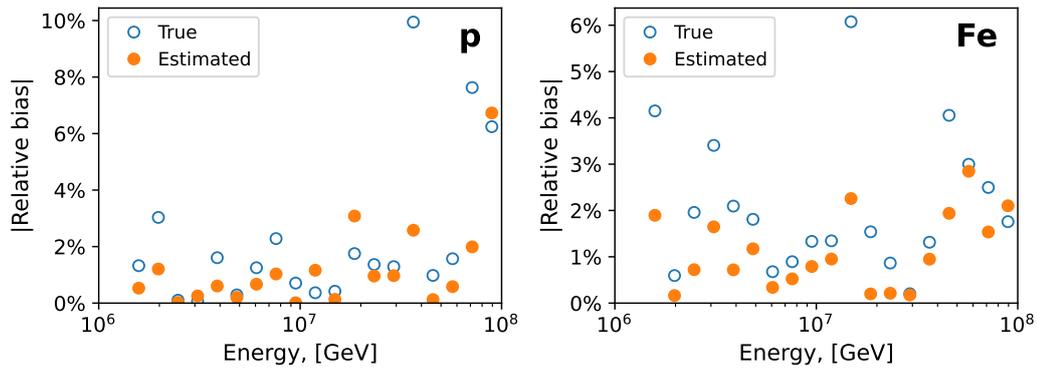


Рисунок 2.20 — Модуль отношения истинного (синие выколотые точки) и оценённого (оранжевые) смещения числа протонов (слева) и железа (справа) к истинным значениям числа протонов и железа в Монте-Карло с генератором QGSJet-II.02 в зависимости от энергии первичной частицы

Из рисунка видно, что значения оценок смещений хотя и не в точности соответствуют истинным значениям, но в целом характер их поведения соответствует истинному, поэтому использование этих оценок в качестве систематической оценки оправдано.

Неопределённости значений в матрицах отклика. Эта неопределённость происходит из-за ограниченной точности матрицы отклика \mathbf{R} и так же как и в предыдущем случае даст соответствующий вклад и в массовый, и в энергетический анфолдинг. Поскольку матрицы отклика, используемые для анфолдинга, извлекаются из Монте-Карло моделирования, объём которого ограничен, значения элементов этих матриц определяются с некоторой статистической погрешностью. Неопределённости элементов \mathbf{R} -матрицы можно определить, принимая количество событий, попавших в ту или иную ячейку матрицы отклика, за пуассоновскую случайную величину и впоследствии пересчитать её в соответствующую погрешность результатов анфолдинга. Пример величины этой неопределённости можно увидеть из Рисунка 2.9, на котором изображены значения диагональных элементов \mathbf{R} -матриц. Из-за большого значения неопределённостей для высоких энергий было решено использовать в массовом анфолдинге объединение матриц по нескольким энергетическим интервалам, как описано в пункте 2.3.2. В пакете `runfold` этот пересчёт неопределённости происходит автоматически в соответствии с формулами байесовского итеративного алгоритма [84; 85].

Наконец, **неопределённость из-за замены при последовательном анфолдинге** в Формуле (2.4). Как видно из формулы, проведённая замена

вероятностей приводит к смещению получаемых результатов относительно точных на величину δ , такую что

$$1 + \delta = \frac{P(A_i^T | E_j^T)}{P(A_i^T | E_j^R)}.$$

Поскольку, как и в случае смещения из-за регуляризации, точное вычисление возможно только при полном знании истинных результатов, то оценим δ следующим образом. Положим полученные после анфолдинга энергетические спектры индивидуальных массовых компонент (dJ_i/dE) за точные, где i пробегает по типу первичных частиц $\{p, He, C, Si, Fe\}$, аппроксимируем их, чтобы нивелировать эффект от возможных выбросов, согласно ломаному степенному закону (broken power law, BPL)

$$dJ_i/dE = J_i^0 \begin{cases} E^{\gamma_i^1}, & \text{если } E < E_i^{br} \\ E^{\gamma_i^2}, & \text{если } E \geq E_i^{br}, \end{cases} \quad (2.7)$$

где E_i^{br} — энергия излома в спектре i -й массовой компоненты, $\gamma_i^{1/2}$ — спектральный индекс до/после излома. И, таким образом, из полученных аппроксимаций вычислим значения $P(A_i^T | E_j^T)$. На Рисунке 2.21 показан пример того, как выглядят эти аппроксимации для открытой части экспериментальных данных после проведения процедуры анфолдинга с Монте-Карло, основанным на генераторе

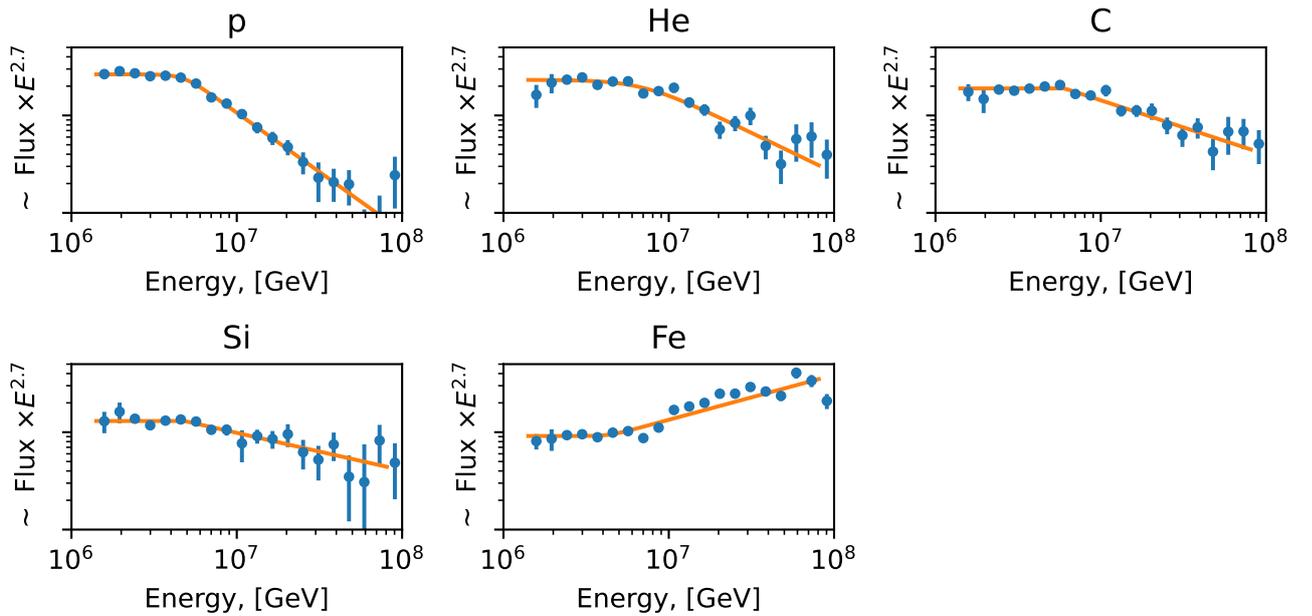


Рисунок 2.21 — Значения потока пяти индивидуальных массовых компонент в зависимости от энергии первичной частицы (синие точки) и их аппроксимация согласно BPL (оранжевые линии)

QGSJet-II.02. Затем сворачиваем этот результат с соответствующей матрицей отклика по энергии, определяя тем самым $P(A_i^T | E_j^R)$. После этого у нас есть всё необходимое для определения δ по Формуле (2.4).

На Рисунке 2.22 показана получившаяся оценка δ для пяти массовых компонент в зависимости от энергии первичной частицы для событий из открытой части экспериментальных данных с Монте-Карло на QGSJet-II.02. Из рисунка видно, что значение погрешности близко к нулю, если доля компоненты не меняется на энергиях порядка разрешения, поскольку в этом случае вероятности в Формуле (2.4) становятся одинаковыми, и, таким образом, средством уменьшения величины этой погрешности в таком подходе является повышение точности реконструкции энергии.

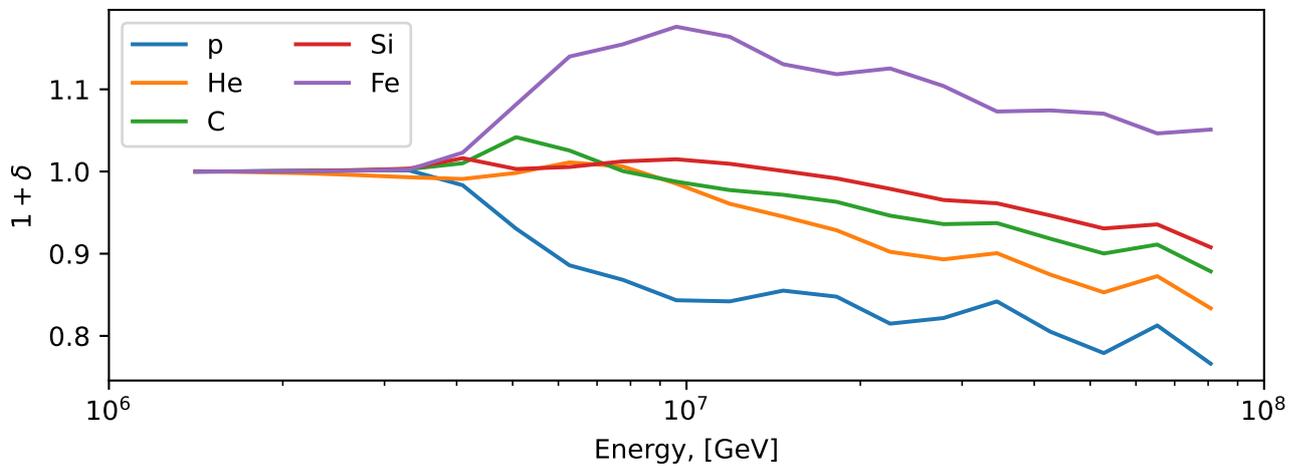


Рисунок 2.22 — Зависимость оценки величины δ для пяти массовых компонент в зависимости от энергии первичной частицы для событий из открытой части экспериментальных данных с Монте-Карло на QGSJet-II.02

2.5.2 Неопределённости Монте-Карло

В эту группу входят неопределённости, связанные с

- массовым составом, заложенным в Монте-Карло моделирование,
- спектральным индексом Монте-Карло моделирования,
- учётом неработающих детекторных станций.

Массовый состав, заложенный в Монте-Карло моделирование. Заложённый состав напрямую влияет на результат энергетического анфолдинга, по-

сколькx энергетическое разрешение различных массовых компонент отличается друг от друга, например, разрешение протонной компоненты ниже чем разрешение железной, как это продемонстрировано на Рисунке 2.23. А значит, используемая в анфолдинге матрица отклика, полученная напрямую из Монте-Карло моделирования заведомо отлична от реальной, если заложенный в него массовый состав отличается от того, который есть в действительности.

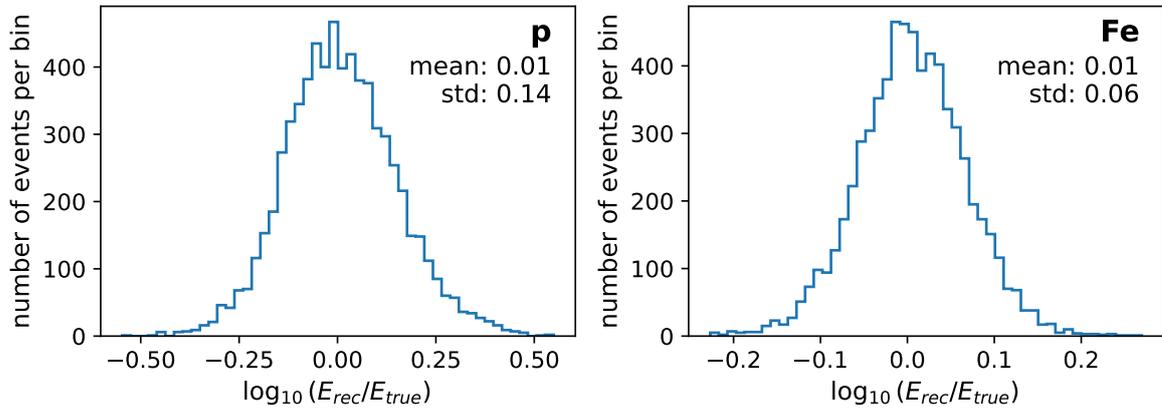


Рисунок 2.23 — Распределение числа событий в Монте-Карло моделировании только протонов (слева) и только железа (справа) с генератором QGSJet-II.02, прошедших используемые критерии отбора, по десятичному логарифму отношения между реконструированной энергией E_{rec} и истинной энергией первичной частицы (E_{true}). Также представлены выборочное среднее (mean), и стандартное отклонение (std)

Идея оценки этой погрешности формулируется так: проведём два дополнительных энергетических анфолдинга с \mathbf{R} -матрицами извлечёнными только из протонной и только из железной компоненты Монте-Карло, сравним результаты со стандартными и используем в качестве консервативной оценки этого типа погрешности отклонение, взвешенное в зависимости от энергии с долями лёгкой ($p + He$) и тяжёлой компоненты (остальные), взятыми из стандартного анфолдинга по массовому типу частицы.

Следующая **неопределённость из-за спектрального индекса Монте-Карло** также влияет только на энергетический анфолдинг. Для определения этой неопределённости перевзвешиваем Монте-Карло к спектральному индексу минус 3, и снова проводим энергетический анфолдинг, сравнивая результаты со стандартными, полученными с индексом минус 2,7. (Изначально, как описано в Разделе 1.2, в Монте-Карло заложен спектральный индекс минус 2 [56],

но для стандартного энергетического анфолдинга оно перевзвешивается к минус 2,7 для большего соответствия экспериментальным данным).

Последняя исследуемая неопределённость возникает **из-за неучёта неработающих детекторов в Монте-Карло**, как обсуждалось в Разделе 2.3.3. Этот фактор влияет как на результат анфолдинга по типу массовой компоненты, так и на энергетический анфолдинг, причём метод учёта этой погрешности для этих двух случаев отличается. В случае энергетического анфолдинга выбирался набор заходов, в которых все детекторы были исправны, и набор заходов с испорченными детекторами. Разница результатов анфолдинга для этих двух наборов данных и объявлялась оценкой этой неопределённости. В случае анфолдинга по типу массовой компоненты, извлекались разницы для соответствующих элементов матриц смешивания, полученных для стандартного и «испорченного» Монте-Карло (подробнее в Разделе 2.3.3). Эти разницы объявлялись ошибками элементов стандартных матриц смешивания (матриц отклика) и стандартными средствами `unfold` распространялись через процедуру анфолдинга (по аналогии с погрешностями элементов матрицы отклика, описанными в Разделе 2.5.1). Итоговая погрешность дополнительно домножалась на коэффициент 0,5, поскольку, как опять же было выяснено в Разделе 2.3.3, около $\sim 50\%$ заходов содержат неработающие детекторы.

2.5.3 Сравнительный анализ неопределённостей

Сначала была проведена проверка процедуры анфолдинга на Монте-Карло моделировании с генератором QGSJet-II.02 и реконструкцией типа первичной частицы с помощью стандартной CNN. Поскольку в моделировании для каждого события указаны истинные значения энергии и типа первичной частицы, то можно оценить согласие между полученными и истинными значениями потока числа частиц в зависимости от энергии. Технически, для этой проверки весь набор Монте-Карло моделирования QGSJet-II.02 разбивался случайным образом на две равные части: первая часть принималась за экспериментальные данные, а вторая — за моделирование, и проводилась описанная выше процедура анфолдинга и процедуры учёта систематических погрешностей этих экспериментальных данных на матрицах отклика из моделирования.

Результаты зависимостей отношений наблюдаемого потока к истинному (заложеному в Монте-Карло моделирование) от энергии первичной частицы без анфолдинга (т. е. потоки получены согласно реконструкции CNN классификатором) и после проведения анфолдинга для индивидуальных массовых компонент и полного потока продемонстрированы на Рисунке 2.24. В случае индивидуальных компонент (p , He , C , Si , Fe) проводился и массовый, и энергетический анфолдинг; в случае полного спектра (all) — только энергетический анфолдинг. Из рисунка видно, что после проведения анфолдинга наблюдаемый поток становится ближе к истинному, а систематические погрешности отражают отклонение от истинных значений, и, таким образом, подтверждается корректность методики.

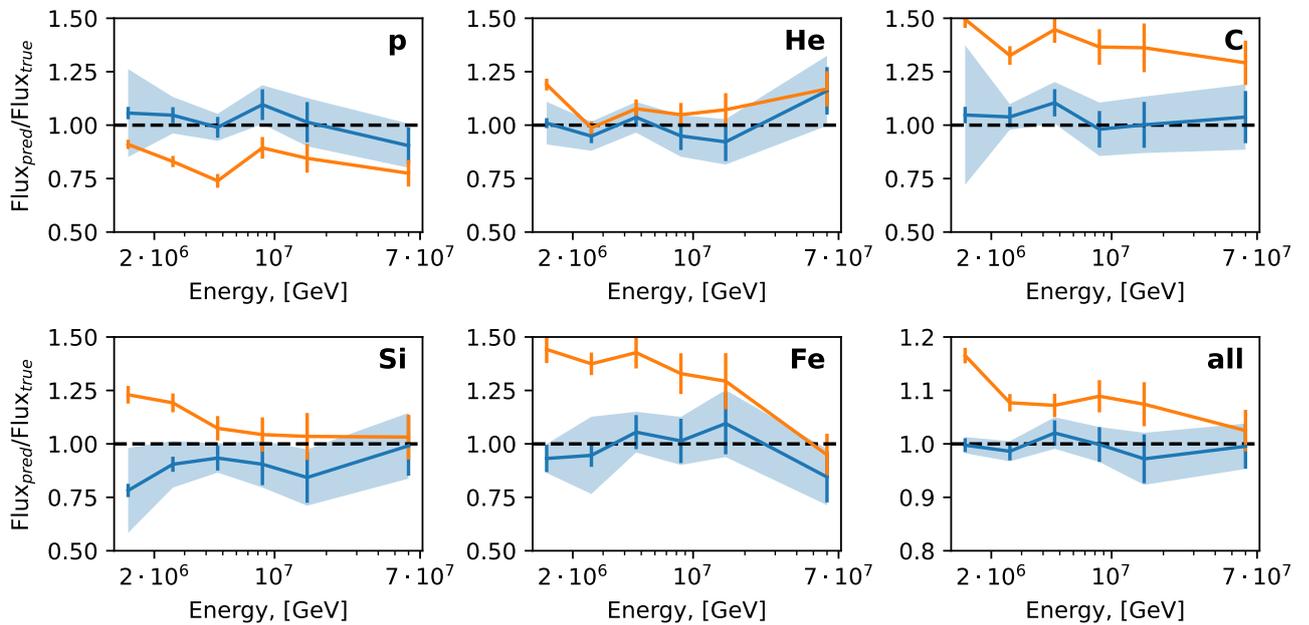


Рисунок 2.24 — Отношение наблюдаемого потока к истинному в зависимости от энергии первичной частицы без процедуры анфолдинга (оранжевый) и после проведения анфолдинга (синий). Вертикальными линиями указаны статистические погрешности, заливкой — систематические. Результаты показаны на примере Монте-Карло моделирования с генератором QGSJet-II.02 и стандартным классификатором CNN

На следующем этапе проводится сравнение исследованных источников неопределённостей на открытой части экспериментальных данных и Монте-Карло с генератором адронных взаимодействий QGSJet-II.02. В данном случае используем QGSJet-II.02, поскольку последние оригинальные результаты KASCADE [23] получены с использованием Монте-Карло основанным на этом

генераторе. Величины неопределённостей также будут оцениваться и для других используемых генераторов, об этом подробнее написано в Разделах 3.1 и 3.2, однако в целом их результаты принципиально не отличаются от полученных для QGSJet-II.02, а методика определения идентична.

Оценки вкладов исследованных источников систематических погрешностей для генератора QGSJet-II.02 приведены в Таблице 2. А именно, диапазоны значений, возникающие в ходе массового (Масс. анф.) и энергетического (Эн. анф.) анфолдинга, которые представляет собой вариацию неопределённости (от минимальной до максимальной) в пределах изучаемого диапазона энергий (1,5–100 ПэВ). Значения неопределённостей для массового анфолдинга были усреднены по всем массовым компонентам. Значения в скобках включают неопределённости, соответствующие выбросам от регуляризации анфолдинга для точек с энергиями выше 50 ПэВ.

Среди источников систематических неопределённостей энергетического анфолдинга выделяется вклад энергетического разрешения, которое вносит значительный вклад в общую неопределённость во всём исследуемом энергетическом диапазоне. Привести к уменьшению этой систематической погрешности в будущих подобных работах может стать использование более сложных алгоритмов (в т. ч. на основе машинного обучения) для реконструкции энергии первичной частицы.

В случае массового анфолдинга наибольший вклад вносит регуляризация анфолдинга и ограниченное Монте-Карло, и если с первым можно бороться улучшением качества обучаемого классификатора (для примера, если рассматривать случай идеального классификатора с единичной матрицей отклика, анфолдинг не нужен вовсе), что представляется маловероятным, поскольку уже в данной работе продемонстрированы результаты реконструкции массовых компонент с помощью различных методов машинного обучения, выходящие на определённый предел, то со вторым можно справиться увеличивая количество данных Монте-Карло моделирования. С учётом развития вычислительных технологий, увеличение объёма Монте-Карло выглядит наиболее простым методом уменьшения систематических погрешностей, и также полезно для обучения моделей машинного обучения (для укрепления их способностей к обобщению и как препятствие к переобучению).

Таблица 2 — Список исследованных систематических неопределённостей и интервал их относительных значений (значения в скобках с учётом выбросов). Использовалось Монте-Карло моделирование с генератором QGSJet-II.02. Суммарная неопределённость («базовая» для QGSJet-II.02) рассчитывается как корень из суммы квадратов всех исследованных неопределённостей

Тип неопределённости	Масс. анф.	Эн. анф.
Регуляризация анфолдинга	1–11(65) %	до 13 %
Ограниченное Монте-Карло	6–20 %	2–15 %
Замена вероятностей	до 13 %	–
Энергетическое разрешение	–	8–10 %
Спектральный индекс	–	0,5–7 %
Неработающие детекторы	1–20 %	4 %
Суммарно: 13–30(79) %	6–24(73) %	11–20 %

2.6 Сравнение с оригинальным результатом KASCADE с генератором QGSJet-II.02

В этом разделе проводится сравнение энергетических спектров (полного и индивидуальных массовых компонент), полученных с помощью методики, описанной в данном исследовании, и последних на текущий момент оригинальных результатов KASCADE [23]. Поскольку последние результаты KASCADE используют Монте-Карло с pre-LHC генератором QGSJet-II.02, то для прямого сравнения в качестве классификатора типа первичной частицы использовалась CNN, обученная на Монте-Карло с генератором адронных взаимодействий QGSJet-II.02 и последующим анфолдингом на матрицах отклика с тем же генератором.

На Рисунке 2.25 показано сравнение полных спектров полученных разработанным в данной работе методом и результатами KASCADE. Полный спектр здесь и далее строится как результат применения к экспериментальным данным только энергетического анфолдинга. В качестве систематических погрешностей текущего результата использованы «базовые» неопределённости с генератором

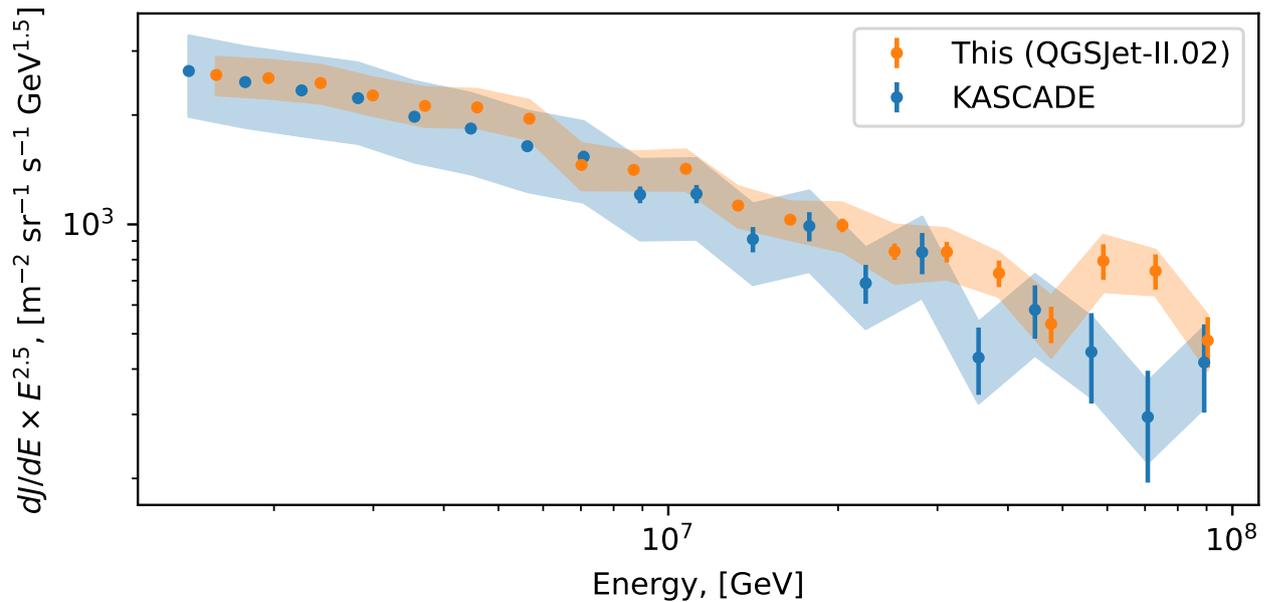


Рисунок 2.25 — Энергетические спектры по всем частицам, полученные в текущей работе на открытой части экспериментальных данных (оранжевый) и в оригинальной работе KASCADE (синий). В обоих случаях результаты основаны на Монте-Карло с генератором QGSJet-II.02. Статистические ошибки изображены в виде вертикальных линий, систематические погрешности — в виде полосы соответствующих цветов

QGSJet-II.02. Из рисунка видно, что результаты хорошо согласуются друг с другом во всём диапазоне исследуемых энергий. Также, на более высоких энергиях для обоих результатов наблюдается волнообразное поведение спектров, что может быть связано с недостатком статистики, и как следствие, флуктуациями в данных и Монте-Карло. По сравнению с оригинальной работой KASCADE, используемый в данной работе подход обеспечивает более низкий уровень систематических неопределённостей, а именно 11–20 % относительно ~25 % для оригинального.

Также необходимо отметить, что статистические погрешности текущего измерения меньше оригинальных, поскольку даже открытая часть экспериментальных данных, на которой проводится сравнение, больше в ~2,5 раза, чем в исследовании KASCADE.

Наконец, было проведено сравнение энергетических спектров пяти индивидуальных массовых компонент. Результаты представлены на Рисунке 2.26. Сравнение с результатами KASCADE прямое, поскольку в обоих случаях использовано Монте-Карло с генератором QGSJet-II.02. Как видно,

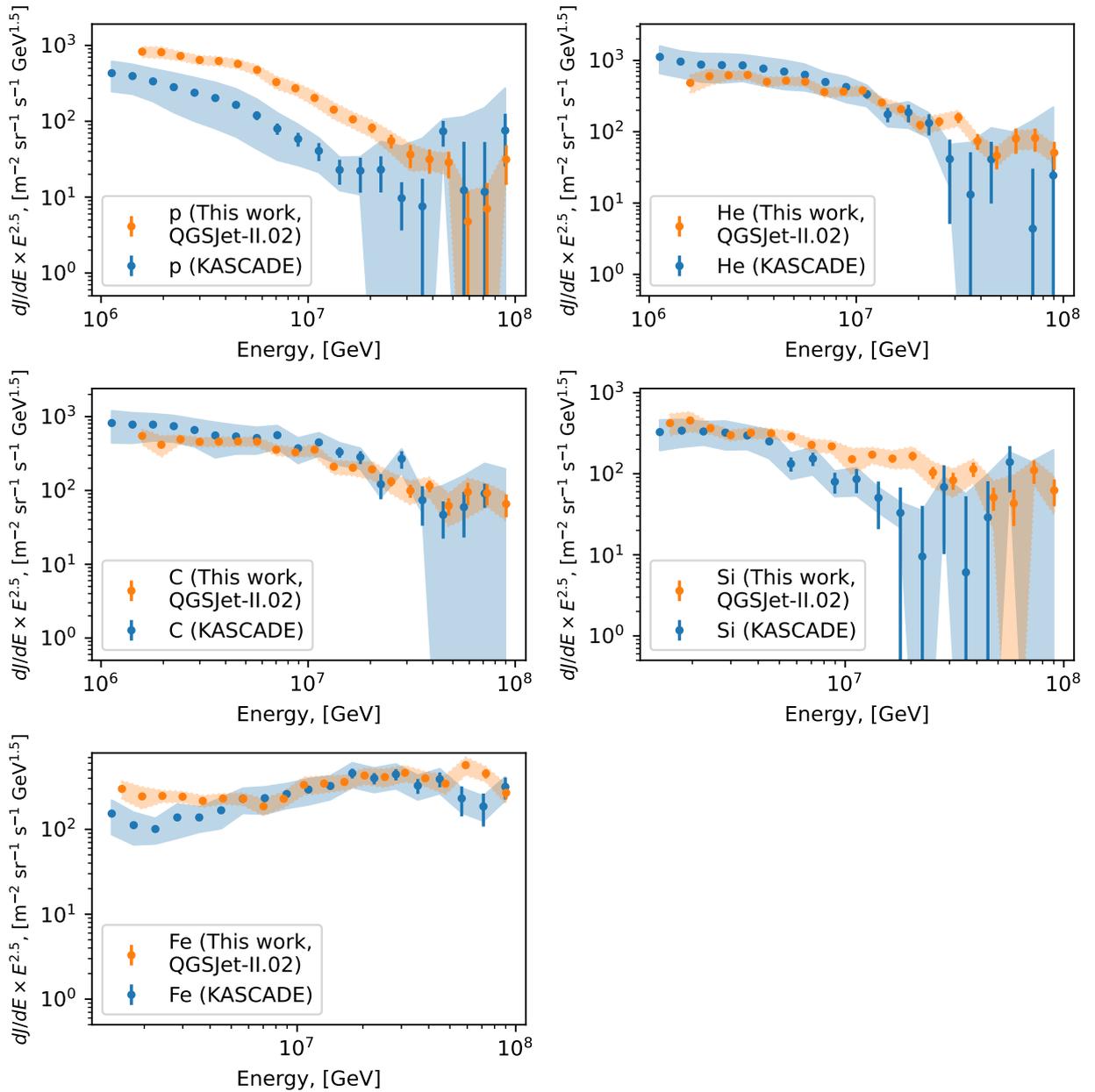


Рисунок 2.26 — Энергетические спектры пяти индивидуальных массовых компонент, полученных в текущей работе на открытой части экспериментальных данных (оранжевый), и оригинальных результатов KASCADE (синий). В обоих случаях результаты основаны на Монте-Карло с генератором QGSJet-II.02. Статистические ошибки изображены в виде вертикальных линий, систематические погрешности — в виде полосы соответствующих цветов

систематические погрешности текущего результата ниже чем в оригинальном методе. А именно, относительные систематические погрешности в энергетическом диапазоне до 10 ПэВ (чтобы исключить вклад точек с малой статистикой) для протонной компоненты составили 13–19 % против 42–64 % для текущего и оригинального метода KASCADE соответственно, для гелиевой компоненты — 13–29 % и 35–44 %, для углерода — 13–32 % и 34–50 %, для кремния — 14–30 % и 32–40 % и, наконец, для железа — 14–29 % и 32–43 %.

Также отмечается заметное различие в спектрах индивидуальной компоненты протонов — текущий метод демонстрирует заметное превышение, например, в около 5 раз при энергии 10 ПэВ и, таким образом, в химическом составе до энергий ~ 5 ПэВ доминируют протоны в отличие от результатов KASCADE, где они изначально подавлены. В то же время, спектры остальных массовых компонент в целом согласуются с измерениями KASCADE с учётом систематических неопределённостей. В продолжение обсуждения можно уже заметить колесообразное поведение в энергетических спектрах протонной и гелиевой компонент, однако детальный анализ индивидуальных спектров будет продемонстрирован в следующей главе, с использованием современного post-LHC генератора адронных взаимодействий QGSJet-II.04.

Как промежуточный итог, можно сказать, что на данном этапе удалось реализовать методику определения индивидуальных энергетических спектров по пяти массовым компонентам с более низкой относительно оригинального исследования KASCADE систематической неопределённостью.

Глава 3. Результаты восстановления спектров массовых компонент космических лучей

В данной главе будет проведён анализ закрытой части экспериментальных данных ($\sim 80\%$ от общего объёма) с использованием Монте-Карло моделирования, основанного на трёх post-LHC адронных генераторах, а именно QGSJet-II.04, EPOS-LHC, Sibyll 2.3c, согласно методике, описанной в Главе 2. Соответственно, будут продемонстрированы, как полный энергетический спектр, так и спектры индивидуальных массовых компонент в изучаемом диапазоне энергий 1,5–100 ПэВ. Также в данной главе будет проведён пересчёт результатов в термины среднего логарифма массы ($\langle \ln A \rangle$), учтена неопределённость, связанная с использованием различных адронных генераторов. Демонстрируется сравнение результатов текущей работы с другими экспериментами: IceTop [24], TALE [25] и LHAASO [29]. Наконец, исследуется поведение полученных спектров массовых компонент и проводится поиск коленоподобной структуры в них, а также проводится обсуждение полученных результатов.

3.1 Восстановление спектров массовых компонент с генератором QGSJet-II.04

Поскольку в предыдущей главе разработанный метод восстановления массового состава был комплексно проверен, то на данном этапе можно переходить к его применению на закрытой части экспериментальных данных с использованием современных (post-LHC) версий генераторов адронных взаимодействий в Монте-Карло. В качестве основной версии используем генератор QGSJet-II.04 для сохранения преемственности с оригинальным исследованием KASCADE, использовавшим Монте-Карло с QGSJet-II.02.

Для восстановления массового состава на закрытой части данных в качестве классификатора массовой компоненты первичной частицы использовалась описываемая в Главе 2 CNN, обученная на тренировочной части Монте-Карло с генератором QGSJet-II.04. Затем производился полный анфолдинг полученных результатов по предсказанным классификатором массовым компонентам и по ре-

конструированной стандартным методом KASCADE энергии первичной частицы согласно матрицам отклика, полученным из тестовой части Монте-Карло (также с генератором QGSJet-II.04).

Систематические погрешности определялись по той же методике, описанной в Разделе 2.5. В Таблице 3 представлены результаты оценки соответствующих погрешностей для генератора QGSJet-II.04. Среди источников систематических погрешностей, значительное влияние во всём исследуемом диапазоне энергий, как и в случае с генератором QGSJet-II.02, оказывает энергетическое разрешение. Систематическая неопределённость от энергетического разрешения оказалась несколько больше чем для QGSJet-II.02, что может быть связано с тем, что оригинальная реконструкция энергии оптимизирована с учётом Монте-Карло QGSJet-II.02 (см. Раздел 1.2).

Энергетические спектры пяти индивидуальных массовых компонент (p , He , C , Si , Fe), полученные на закрытой части экспериментального набора данных с использованием Монте-Карло на QGSJet-II.04 показаны на Рисунке 3.1. Качественное поведение спектров не отличается от тех, что получены с генератором QGSJet-II.02, а именно на низких энергиях доли массовых компонент

Таблица 3 — Список исследованных систематических неопределённостей и интервал их относительных значений с использованием Монте-Карло моделирования с генератором QGSJet-II.04. Суммарная неопределённость («базовая» для QGSJet-II.04) рассчитывается как корень из суммы квадратов всех исследованных неопределённостей

Тип	Масс. анф.	Эн. анф.
Регуляризация анфолдинга	1–23 %	до 10 %
Ограниченное Монте-Карло	5–25 %	2–18 %
Замена вероятностей	до 8 %	–
Энергетическое разрешение	–	13–16 %
Спектральный индекс	–	до 4 %
Неработающие детекторы	2–17 %	4 %
Суммарно: 19–42 %	12–40 %	14–28 %

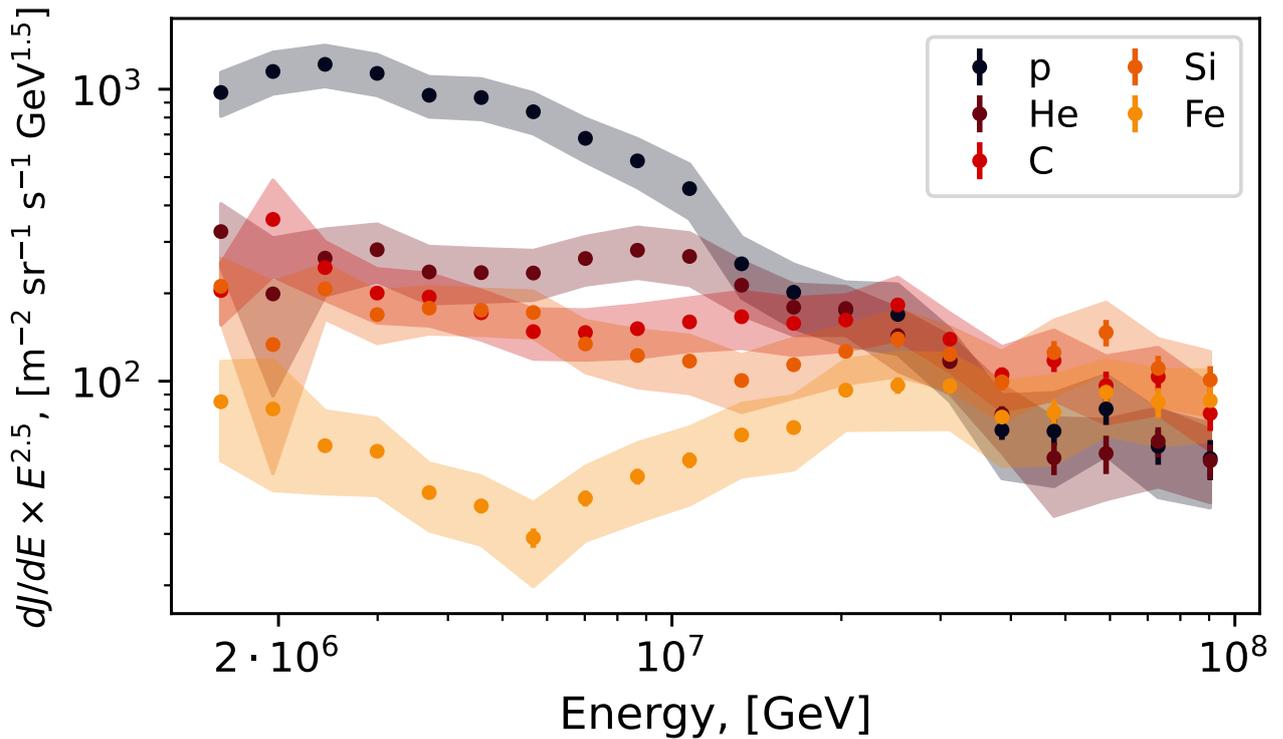


Рисунок 3.1 — Энергетические спектры индивидуальных массовых компонент, полученные в данном исследовании на закрытой части экспериментальных данных, с использованием Монте-Карло, основанном на QGSJet-II.04. Статистические ошибки изображены в виде вертикальных линий, «базовые» систематические погрешности — в виде полос соответствующих цветов

убывают в порядке увеличения массового числа, и изменение этого поведения происходит на энергиях 20–50 ПэВ. Спектры индивидуальных компонент ведут себя достаточно плавно за исключением второй энергетической точки ~ 2 ПэВ, где виден скачок в спектрах *He*, *C*, *Si* массовых компонент, связанный с проведением анфолдинга по массовому типу (в связи с этим «базовая» систематическая неопределённость в этой точке заметно выше чем в соседних). Однако так же как и в случае спектров, полученных на открытой части данных с генератором QGSJet-II.02, они отличаются от оригинальных KASCADE особенно для протонной компоненты, которая в данном случае доминирует на низких энергиях.

3.2 Учёт неопределённостей, связанных с различными генераторами адронных взаимодействий

Для того чтобы учесть неопределённость, связанную с использованием генераторов из различных семейств (QGSJet-II.04, EPOS-LHC, Sibyll 2.3c), в дополнение к уже проведённому анфолдингу с матрицами отклика, извлечёнными из Монте-Карло с генератором QGSJet-II.04, выполняется полный анфолдинг с помощью двух остальных наборов Монте-Карло. Для этого во всех трёх случаях использовалась реконструкция массовой компоненты отдельных событий с помощью CNN, обученной на QGSJet-II.04. Таким образом, различные семейства генераторов адронных взаимодействий оказывают влияние на соответствующие матрицы отклика, в данном случае делая их менее диагональными (подробнее см. раздел 2.3.4), что сказывается на качестве проводимого анфолдинга. Однако это всё же более предпочтительный вариант чем, например, обучение классификаторов типов первичной частицы для каждого набора Монте-Карло, поскольку, как было отмечено ранее (см. Раздел 1.2), объёмы Монте-Карло EPOS-LHC и Sibyll 2.3c меньше чем в случае QGSJet-II.04, а также сам процесс обучения классификаторов требует дополнительного времени и тщательной валидации в каждом случае.

Для каждого генератора определялась полоса соответствующей «базовой» систематической неопределённости. Полученные энергетические спектры индивидуальных массовых компонент и полный спектр для генераторов QGSJet-II.04, EPOS-LHC, Sibyll 2.3c продемонстрированы на Рисунке 3.2. Из представленного рисунка видно, что в целом поведение энергетических спектров индивидуальных массовых компонент и полного спектра для всех трёх генераторов согласуется, однако нужно отметить выбросы в спектрах гелиевой и железной компонент, полученные с генератором EPOS-LHC. В случае гелия в области энергий ~ 5 ПэВ, выбросы обусловлены систематической погрешностью в процедуре анфолдинга, которая составила 100%. Такая погрешность возникает из-за достаточно высокой величины смешивания между протонами и гелием для EPOS-LHC (что можно увидеть из Рисунка 2.16, (б)) и доминирования протонов в этой области, из-за чего точное определение гелиевой компоненты затруднено. Рассматривая спектр железной компоненты, в случае EPOS-LHC можно отметить резкое поведение, которое также связано с процедурой анфолдинга спектров по типу первичной

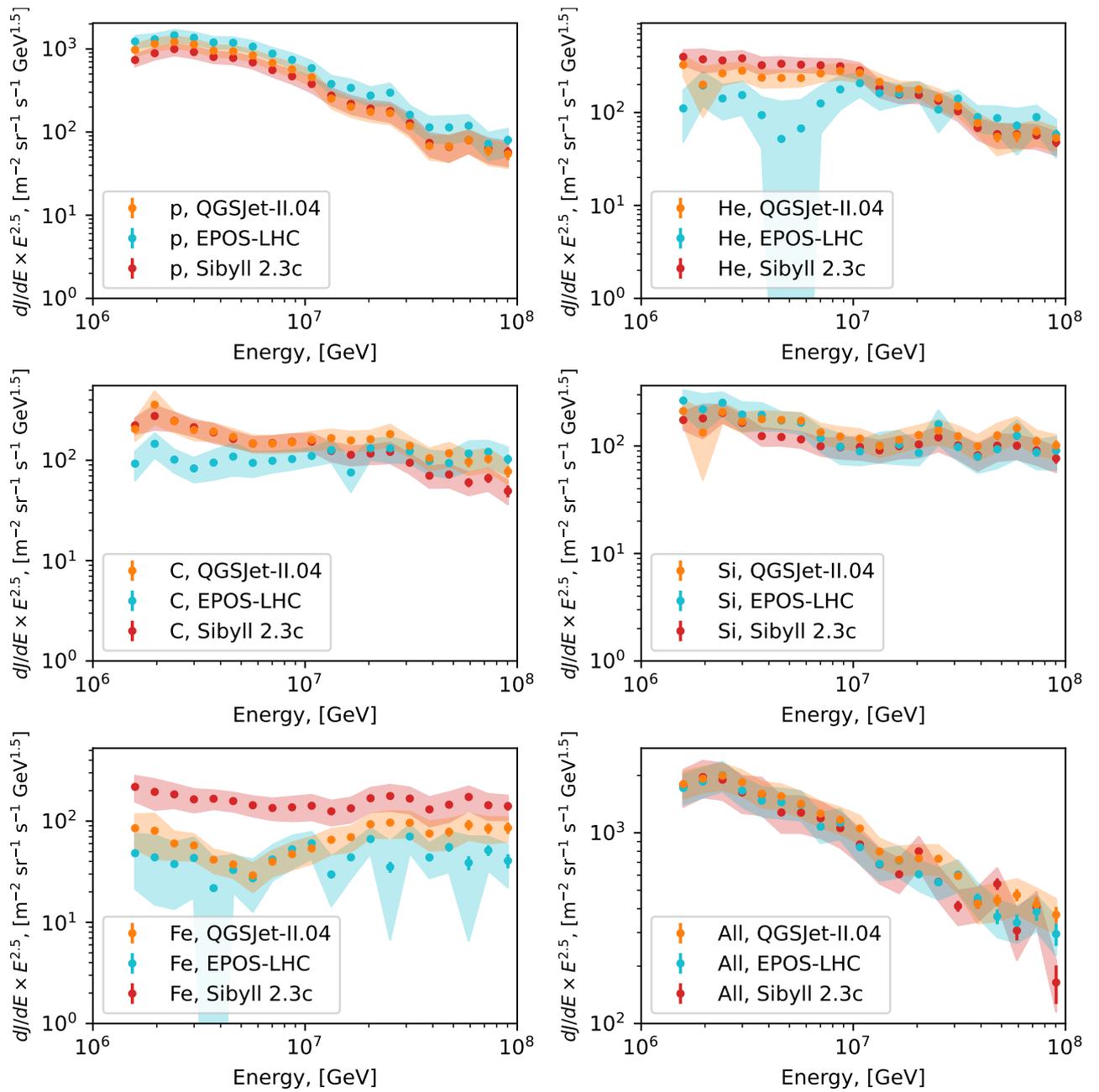


Рисунок 3.2 — Энергетические спектры индивидуальных массовых компонент (p , He , C , Si , Fe) и полный спектр, полученные в данном исследовании на закрытой части экспериментальных данных, для трёх генераторов адронных взаимодействий: QGSJet-II.04 (оранжевый), EPOS-LHC (бирюзовый), Sibyll 2.3c (красный). Статистические ошибки изображены в виде вертикальных линий, «базовые» систематические погрешности для каждого генератора — в виде полос соответствующих цветов

частицы, поскольку, как было отмечено ранее в Разделе 2.3.4, EPOS-LHC среди двух других генераторов наиболее «протонный» и относит к железу меньшую долю событий, чем остальные, что в дополнение к и так малой доле железа в массовом составе увеличивает систематическую неопределённость анфолдинга. По аналогии с EPOS-LHC объясняется поведение спектра железной компоненты, полученной для Sibyll 2.3с. Поскольку, как было отмечено, Sibyll 2.3с является наиболее «железным» генератором относительно остальных, то в основном на нём Монте-Карло к железу причисляются события, характеристики которых соответствуют более лёгким ядрам в остальных генераторах. Таким образом, наблюдается более высокая доля железной компоненты для Sibyll 2.3с относительно других генераторов.

Далее необходимо было оценить неопределённость, связанную с использованием различных генераторов адронных взаимодействий, которая в данном исследовании будет называться теоретической неопределённостью. Теоретическая неопределённость конкретного энергетического спектра определялась как диапазон между минимальной и максимальной границей среди «базовых» систематических погрешностей различных моделей для соответствующей энергии в каждом энергетическом бине. Метод построения границ теоретической неопределённости на примере полных энергетических спектров показан на Рисунке 3.3 чёрными линиями, которые, как видно из рисунка, охватывают диапазон «базовых» неопределённостей трёх современных (post-LHC) генераторов (QGSJet-II.04, EPOS-LHC, Sibyll 2.3с). Однако необходимо отметить, что теоретические неопределённости лишь демонстрируют текущую степень расхождения между этими тремя генераторами адронных взаимодействий, и, следовательно, полученные результаты в границах теоретических неопределённостей не нужно толковать в смысле модельно независимых, поскольку все три исследованных генератора могут в будущем оказаться некорректными и истинные значения, таким образом, не будут покрыты теоретическими неопределённостями.

Проведя аналогичную процедуру для индивидуальных массовых компонент, были определены их теоретические неопределённости. На Рисунке 3.4 показаны энергетические спектры индивидуальных массовых компонент с учётом теоретических неопределённостей. Таблицы 4 — 9 с полученными результатами представлены в Приложении А. Дополнительно, в иллюстративных целях, на Рисунке 3.4 также показаны оригинальные результаты KASCADE [23], (с генератором QGSJet-II.02 и без учёта влияния различных генераторов).

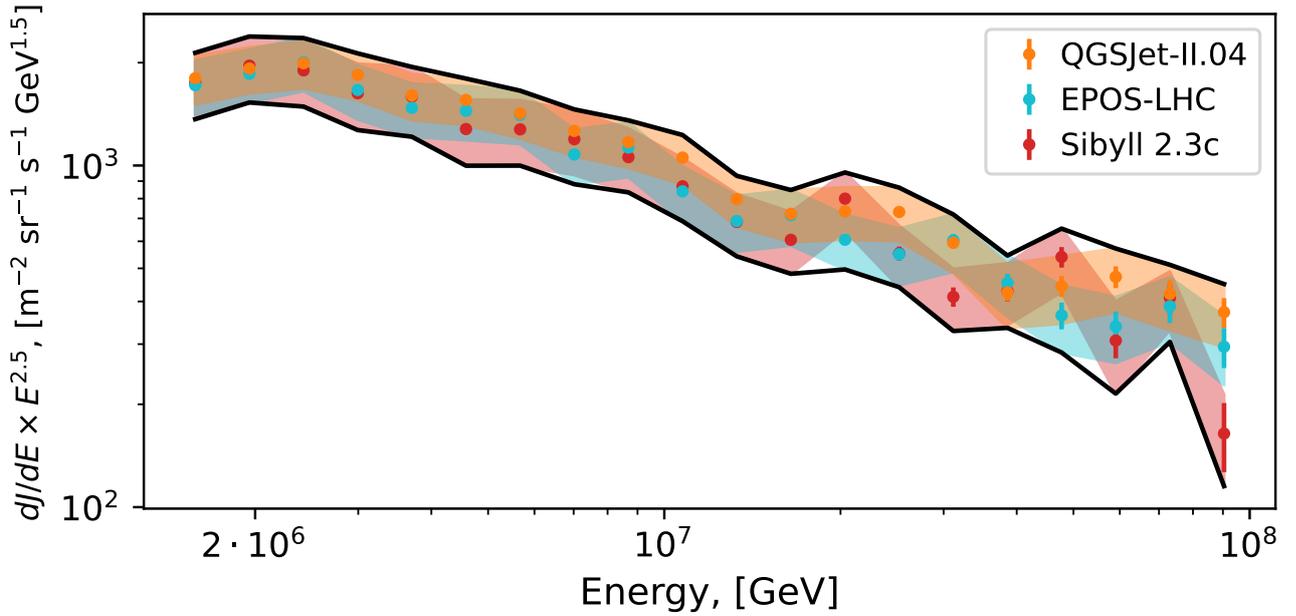


Рисунок 3.3 — Полные энергетические спектры, полученные на закрытой части экспериментальных данных для трёх генераторов адронных взаимодействий: QGSJet-II.04 (оранжевый), EPOS-LHC (бирюзовый), Sibyll 2.3c (красный). Точками обозначены центральные значения, вертикальными линиями — статистические погрешности, полосами — «базовые» систематические погрешности для соответствующего генератора. Чёрными линиями показаны верхняя и нижняя границы теоретической неопределённости

Из рисунка можно отметить, что полные энергетические спектры согласуются с оригинальными результатами KASCADE с учётом неопределённостей для всех исследованных генераторов адронных взаимодействий, хотя также видно например, что полученные центральные точки (соответствующие QGSJet-II.04) лежат систематически ниже оригинальных.

Также из представленного рисунка видно, что протонная компонента на низких энергиях доминирует для всех трёх исследованных генераторов адронных взаимодействий и, соответственно, во всех случаях демонстрирует превышение над оригинальными результатами KASCADE. Причём наблюдаемый на Рисунке 3.4 поток протонной компоненты даже выше чем полученный в данной работе ранее на закрытой части данных с генератором QGSJet-II.02 (см. Рисунок 2.26), например, в ~ 2 раза при энергии 10 ПэВ.

Исключая выбросы, полученные при анфолдинге с генератором EPOS-LHC, а также спектр железной компоненты Sibyll 2.3c, о которых упоминалось в предыдущем пункте, индивидуальные спектры, полученные с помощью

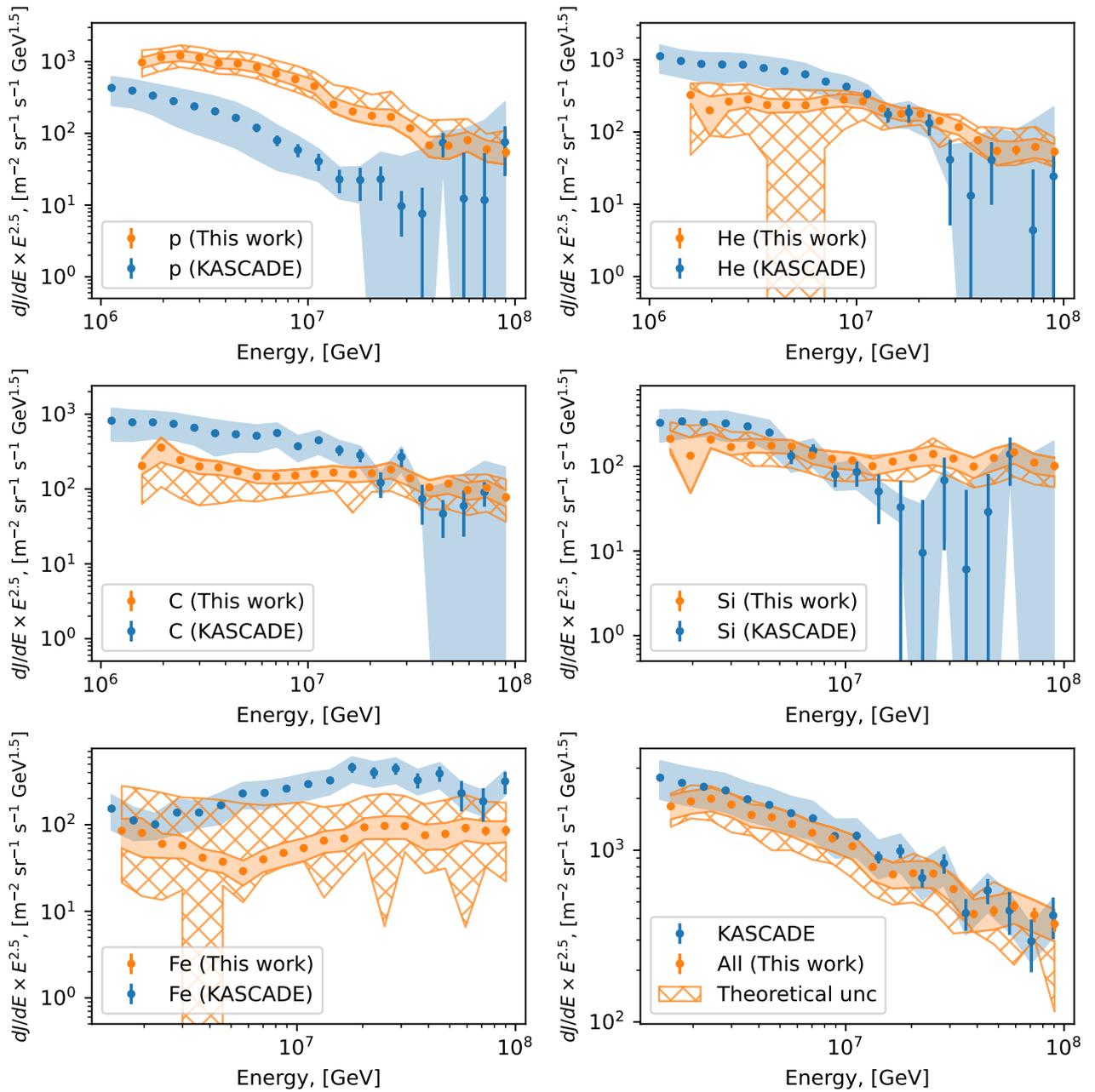


Рисунок 3.4 — Энергетические спектры индивидуальных массовых компонент с учётом теоретической неопределённости, полученные в данном исследовании на закрытой части экспериментальных данных. Для иллюстрации продемонстрированы оригинальные спектры KASCADE (синим), полученные на Монте-Карло с генератором QGSJet-II.02. Статистические ошибки изображены в виде вертикальных линий, систематические погрешности — в виде полос соответствующих цветов, неопределённости с учётом теоретической обозначены штриховкой. Систематические погрешности спектров данной работы, показанные на рисунке, являются «базовыми» для генератора QGSJet-II.04

трёх post-LHC адронных генераторов, качественно демонстрируют похожее поведение в зависимости от энергии. Это может говорить о том, что новые версии генераторов адронных взаимодействий становятся более согласованными, поэтому переобработка архивных данных подобных KASCADE экспериментов с их использованием является важной задачей.

3.3 Представление спектров в виде среднего логарифма массы

Некоторые эксперименты, например TALE [25], представляют свои результаты в виде зависимости среднего массового логарифма ($\langle \ln A \rangle$) от энергии первичной частицы. Для сравнения с ними полезно пересчитать результаты текущего исследования в этих терминах. В данном случае, воспользуемся следующей формулой

$$\langle \ln A \rangle (E) = \sum_{i=1}^5 f_i(E) \cdot \ln A_i, \quad (3.1)$$

где $f_i(E)$ — доля i -ой массовой компоненты ($\{p, He, C, Si, Fe\}$) в точке с энергией E , A_i — массовое число i -ой компоненты.

В результате были получены зависимости, изображённые на Рисунке 3.5. Численные значения представлены в Таблице 10 в Приложении А. Как видно из рисунка, характер зависимости $\langle \ln A \rangle$ от энергии первичной частицы согласуется для всех трёх рассмотренных генераторов адронных взаимодействий (QGSJet-II.04, EPOS-LHC, Sibyll 2.3c), а именно, начиная с энергий ~ 10 ПэВ, наблюдается заметный рост $\langle \ln A \rangle$, однако присутствует заметная разница в абсолютных значениях величин. Это, в том числе отражается на теоретической неопределённости, которая в результате соизмерима по величине с «базовыми» неопределённостями. Также из рисунка можно отметить, что массовый состав соответствующий генератору EPOS-LHC оказался систематически более лёгким, а в случае Sibyll 2.3c — более тяжёлым, что соответствует свойствам матриц смешивания, описанных в разделе 2.3.4.

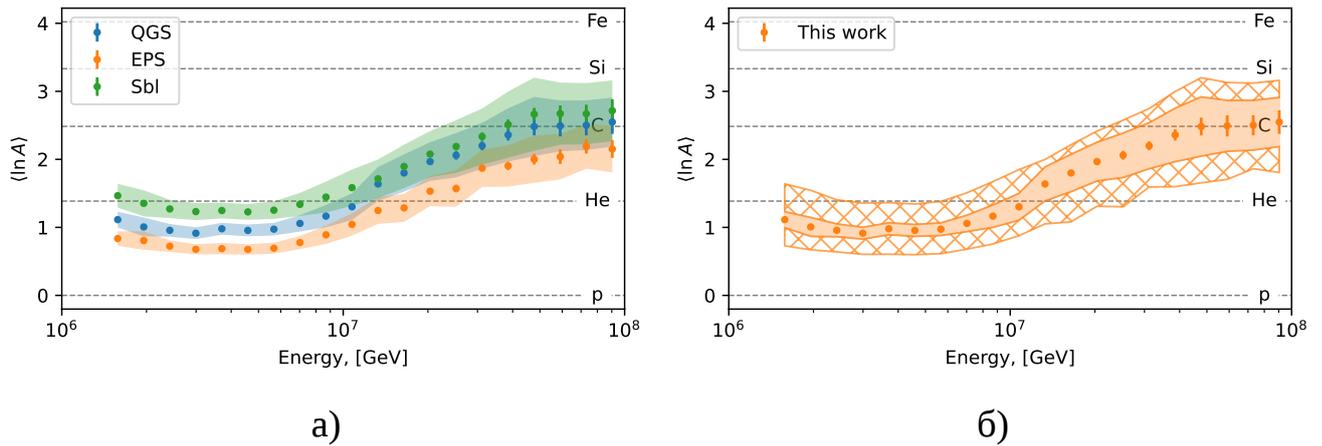


Рисунок 3.5 — Зависимость среднего логарифма массы ($\langle \ln A \rangle$) от энергии первичной частицы. На рисунке (а) представлены полученные результаты с использованием трёх различных генераторов адронных взаимодействий QGSJet-II.04 (синий), EPOS-LHC (оранжевый), Sibyll 2.3c (зелёный), соответствующими полосами выделены «базовые» неопределённости для каждого генератора. На рисунке (б) представлен финальный результат с учётом теоретической неопределённости, обозначенной штриховкой, центральные значения и статистические погрешности соответствуют генератору QGSJet-II.04, «базовые» погрешности QGSJet-II.04 показаны полосой

3.4 Сравнение с результатами других экспериментов

Интересно и важно для понимания места текущего исследования в общем контексте сравнить полученные в данной работе результаты с результатами других экспериментов в изучаемом энергетическом диапазоне 1,5–100 ПэВ, а именно IceTop [24], TALE [25], LHAASO [29].

В первую очередь проведём сравнение спектров индивидуальных массовых компонент. Среди представленных выше экспериментов, на текущий момент лишь IceTop опубликовал покомпонентные энергетические спектры, поэтому в данном случае используем только их результаты.

Сравнение энергетических спектров индивидуальных массовых компонент, полученных в данной работе, с IceTop представлено на Рисунке 3.6. Нужно отметить, что IceTop выделяет не пять, а только четыре массовые компоненты (p , He , O , Fe). А также IceTop использует в качестве генератора адронных взаимодействий pre-LHC модель Sibyll 2.1 (авторы в оригинальной работе [24] оговариваются, что ими исследовались различные генераторы, но в финальных

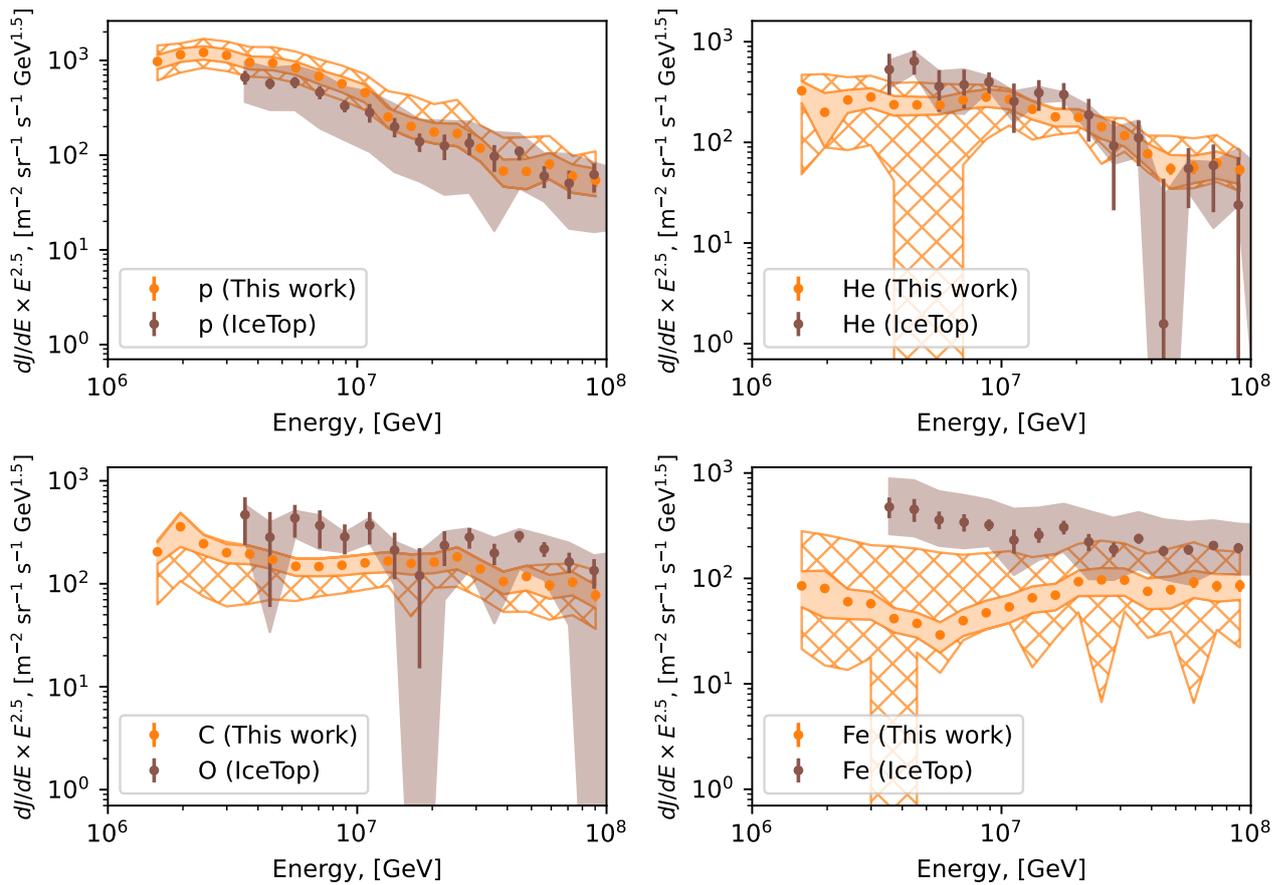


Рисунок 3.6 — Энергетические спектры индивидуальных массовых компонент, полученные в данном исследовании на закрытой части экспериментальных данных с учётом теоретической неопределённости, и результаты IceTop [24] (использует Sibyll 2.1). Статистические ошибки изображены в виде вертикальных линий, систематические погрешности — в виде полос соответствующих цветов неопределённости с учётом теоретической обозначены штриховкой. Систематические погрешности спектров данной работы, показанные на рисунке, являются «базовыми» для генератора QGSJet-II.04

результатах учёта вклада различных генераторов нет), поэтому сравнение не вполне строгое. Однако даже в таком случае видно, что с учётом теоретических неопределённостей результаты, полученные в текущем исследовании хорошо согласуются с IceTop (в том числе и протонная компонента, избыток которой по сравнению с оригинальным KASCADE обсуждался выше). Кроме того, интересно, что результаты, полученные в данной работе с генератором Sibyll 2.3с лежат ближе к результатам IceTop (полученным на Sibyll 2.1), в чём можно убедиться, возвратившись к Рисунку 3.2. Особенно показательное это проявляется в случае протонного и железного спектров.

Наконец, важно подчеркнуть, что как видно из приведённого рисунка, «базовые» систематические погрешности спектров текущей работы ниже чем в исследовании IceTop, например относительные неопределённости для спектра протонов в зависимости от энергии составляют 16–35 % в текущей работе (для QGSJet-II.04), 45–80 % в IceTop. Причём даже с учётом теоретических неопределённостей точность полученных в данной работе индивидуальных спектров сравнима с IceTop (в которые не включена подобная неопределённость). Это ещё раз подчёркивает важность идеи повторной обработки архивных данных с использованием новых методов анализа.

Как было отмечено ранее, часто эксперименты в качестве результата для массового состава приводят зависимость среднего логарифма массового числа от энергии первичной частицы. Поскольку результаты текущей работы представлены и в таком виде (см. пункт 3.3), то оказывается возможным сравнить их с другими экспериментами. Таким образом, сравнение зависимостей $\langle \ln A \rangle$ от энергии первичной частицы продемонстрировано на Рисунке 3.7. На нём показаны результаты текущего исследования, а также результаты экспериментов IceTop (с использованием генератора Sibyll 2.1) и TALE (EPOS-LHC) и LHAASO [29] (с использованием генераторов QGSJet-II.04, EPOS-LHC, Sibyll 2.3d).

Из верхнего рисунка видно, что текущие результаты согласуются с двумя другими (TALE, IceTop) в рамках теоретических неопределённостей, более того, текущий результат с использованием генератора EPOS-LHC тяготеет к результатам TALE а, как было отмечено выше в индивидуальных спектрах и подтверждается в данном случае, Sibyll 2.3c — к результатам IceTop. Кроме того, «базовые» неопределённости в текущей работе ниже соответствующих в TALE и IceTop.

Из сравнения текущих результатов с экспериментом LHAASO на нижнем рисунке можно также отметить хороший уровень согласия между ними с учётом теоретических неопределённостей, в то время как сам уровень неопределённостей LHAASO заметно ниже, чем в текущей работе.

В целом, качественное поведение $\langle \ln A \rangle$ в зависимости от энергии первичной частицы согласуется для всех представленных результатов, а именно в исследуемом диапазоне энергий 1,5–100 ПэВ наблюдается резкий переход к более тяжёлому массовому составу. Однако в абсолютных числах наблюдается заметное расхождение между различными экспериментами, которое частично объясняется использованием различных генераторов адронных взаимодействий.

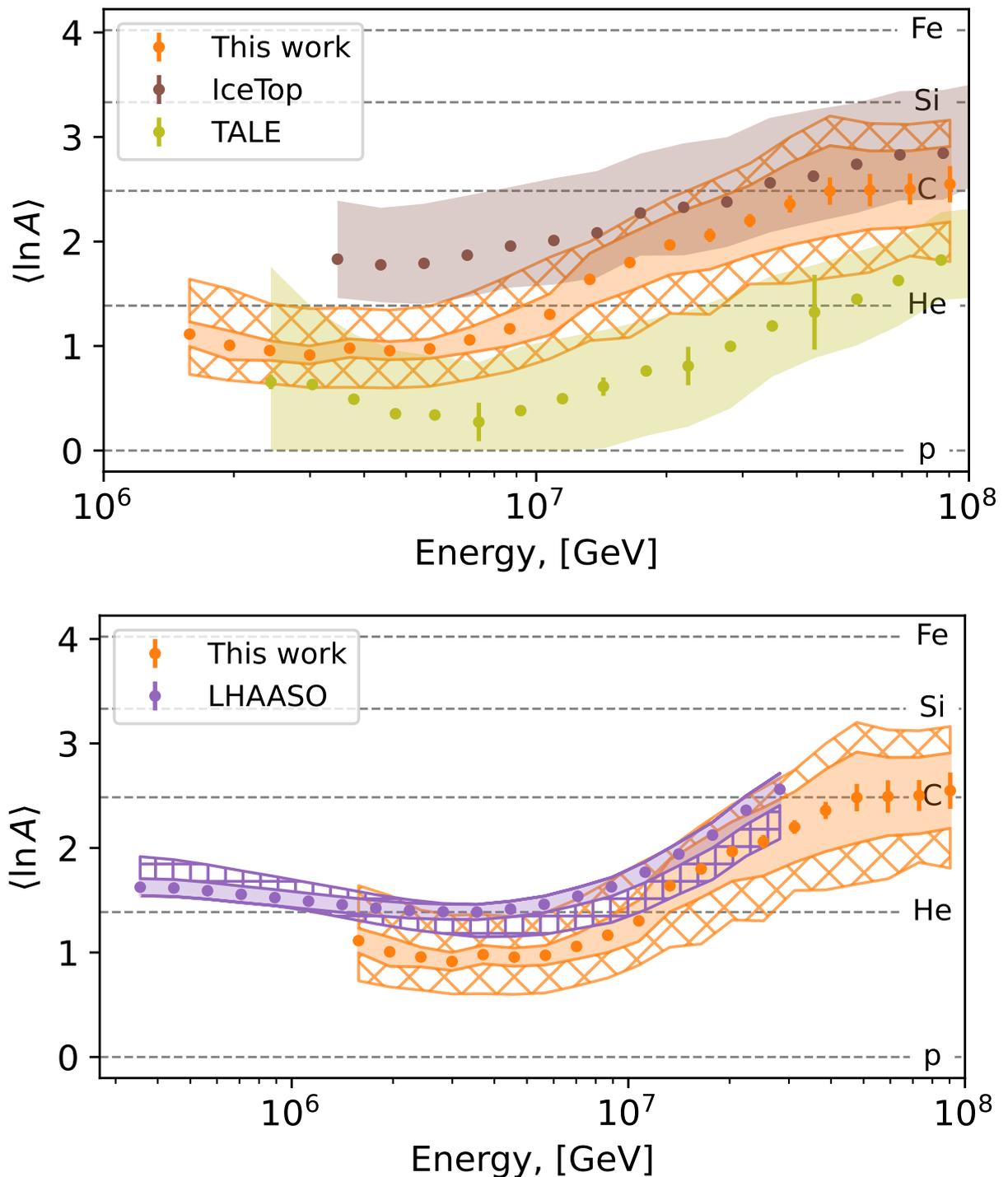


Рисунок 3.7 — Зависимость среднего массового логарифма ($\langle \ln A \rangle$) от энергии первичной частицы для текущего метода (оранжевый, QGSJet-II.04), результатов IceTop [24] (коричневый, Sibyll 2.1) и TALE [25] (оливковый, EPOS-LHC) а также LHAASO [29] (фиолетовый, QGSJet-II.04). Сплошной заливкой показаны «базовые» систематические погрешности соответствующих генераторов адронных взаимодействий, а штриховкой теоретические неопределённости. Теоретические неопределённости в текущей работе определялись с использованием генераторов QGSJet-II.04, EPOS-LHC, Sibyll 2.3c для LHAASO — QGSJet-II.04, EPOS-LHC, Sibyll 2.3d

Несоответствия между генераторами, проявляющиеся в теоретических неопределенностях как результатов данной работы так и LHAASO, всё ещё являются серьёзным фактором систематической неточности.

3.5 Поиск коленоподобных структур в спектрах индивидуальных компонент

Отдельный важный вопрос заключается в изучении поведения энергетических спектров индивидуальных массовых компонент, поиске в них изломов и коленоподобных структур.

В данной работе используется следующая процедура поиска. На первом этапе каждый индивидуальный и полный спектры (dJ/dE) аппроксимируются двумя функциями: степенной функцией (power law — PL),

$$dJ/dE = J_0 \cdot E^\gamma, \quad (3.2)$$

где J_0 — нормализационный коэффициент, γ — спектральный индекс; а также ломаной степенной функцией (broken power law — BPL),

$$dJ/dE = J_0 \cdot \begin{cases} (E/E_{br})^{\gamma_1}, & \text{если } E < E_{br}, \\ (E/E_{br})^{\gamma_2}, & \text{иначе,} \end{cases} \quad (3.3)$$

где $\gamma_{1/2}$ — спектральные индексы до и после излома при энергии E_{br} .

Аппроксимации в каждом случае минимизировали χ^2 -функцию для соответствующего спектра:

$$\chi^2 = \sum_i \frac{(y_i - f(x_i))^2}{\sigma_{i,stat}^2 + \sigma_{i,bas\ sys}^2}, \quad (3.4)$$

где y_i — измеренное значение в точке x_i , $f(x_i)$ — значение функции аппроксимации в x_i , $\sigma_{i,stat}$ — статистическая погрешность, $\sigma_{i,bas\ sys}$ — «базовая» систематическая погрешность для QGSJet-II.04 в точке x_i . Аппроксимация проводилась с помощью минимизатора Minuit [87] в библиотеке iminuit.

Для определения статистической значимости, насколько предпочтительнее использовать гипотезу BPL относительно PL для каждого спектра, применялся тест отношения правдоподобия, в котором используется факт, что величина $\Delta\lambda = \chi_{PL}^2 - \chi_{BPL}^2$ — улучшение качества аппроксимации при использовании

«длинной» гипотезы ВРЛ относительно использования «короткой» гипотезы РЛ, распределена согласно закону хи-квадрат с q степенями свободы, где q — число ограничений на параметры в «короткой» гипотезе (относительно «длинной»), в данном случае два. Таким образом, рассчитывалось p -значение и переводилось в термины значимости (σ) как

$$\sigma = \Phi^{-1}(1 - p/2), \quad (3.5)$$

где $\Phi(x)$ — функция распределения $\mathcal{N}(0, 1)$. В этих терминах p -значение 0,05 соответствует 2σ значимости.

Результаты аппроксимаций энергетических спектров и численные значения оптимизированных параметров показаны на Рисунке 3.8.

Среди показанных спектров, статистически значимый излом удалось обнаружить для протонной и для гелиевой массовых компонент. В случае p , использование ВРЛ приводит к $\Delta\lambda = 30,9$ взамен на две степени свободы, и, таким образом к p -значению $1,96 \cdot 10^{-7}$ ($5,2 \sigma$), энергия колена согласно аппроксимации составляет $4,4 \pm 0,7$ ПэВ. Для He компоненты, $\Delta\lambda = 18,2$ приводит к p -значению $1,07 \cdot 10^{-4}$ ($3,9 \sigma$) для энергии колена $11,4 \pm 2,8$ ПэВ.

Также в случае спектра массовой компоненты железа впервые наблюдается указание на излом ($\Delta\lambda = 8,2$) с p -значением $1,62 \cdot 10^{-2}$ ($2,4 \sigma$) при энергии $4,5 \pm 1,2$ ПэВ. Однако как видно из Рисунка 3.8, если спектры p и He демонстрируют коленоподобное поведение, то в случае Fe спектральный индекс после излома уменьшается.

Таким образом, отмечается, что хотя коленоподобное поведение в энергетических спектрах протонной и гелиевой компонент наблюдалось в работах [13; 23; 25], в данном исследовании впервые оно подтверждено с высоким уровнем значимости. Также важно подчеркнуть, что полученные из аппроксимации ВРЛ положения энергии колена в спектрах протонов $\sim 4,4$ ПэВ и гелия $\sim 11,4$ ПэВ в пределах погрешностей согласуются с теориями, предполагающими прямую зависимость энергий колена от заряда соответствующего ядра. В то же время, в энергетических спектрах углеродной, кремниевой и железной компоненты не наблюдается каких-либо значимых признаков присутствия подобной коленоподобной структуры. Это может объясняться тем, что если положение энергии колена для первичного ядра действительно пропорционально его заряду, то в спектрах железа оно должно располагаться выше максимальной исследуемой энергии 100 ПэВ, поэтому не может наблюдаться здесь; в случае кремния

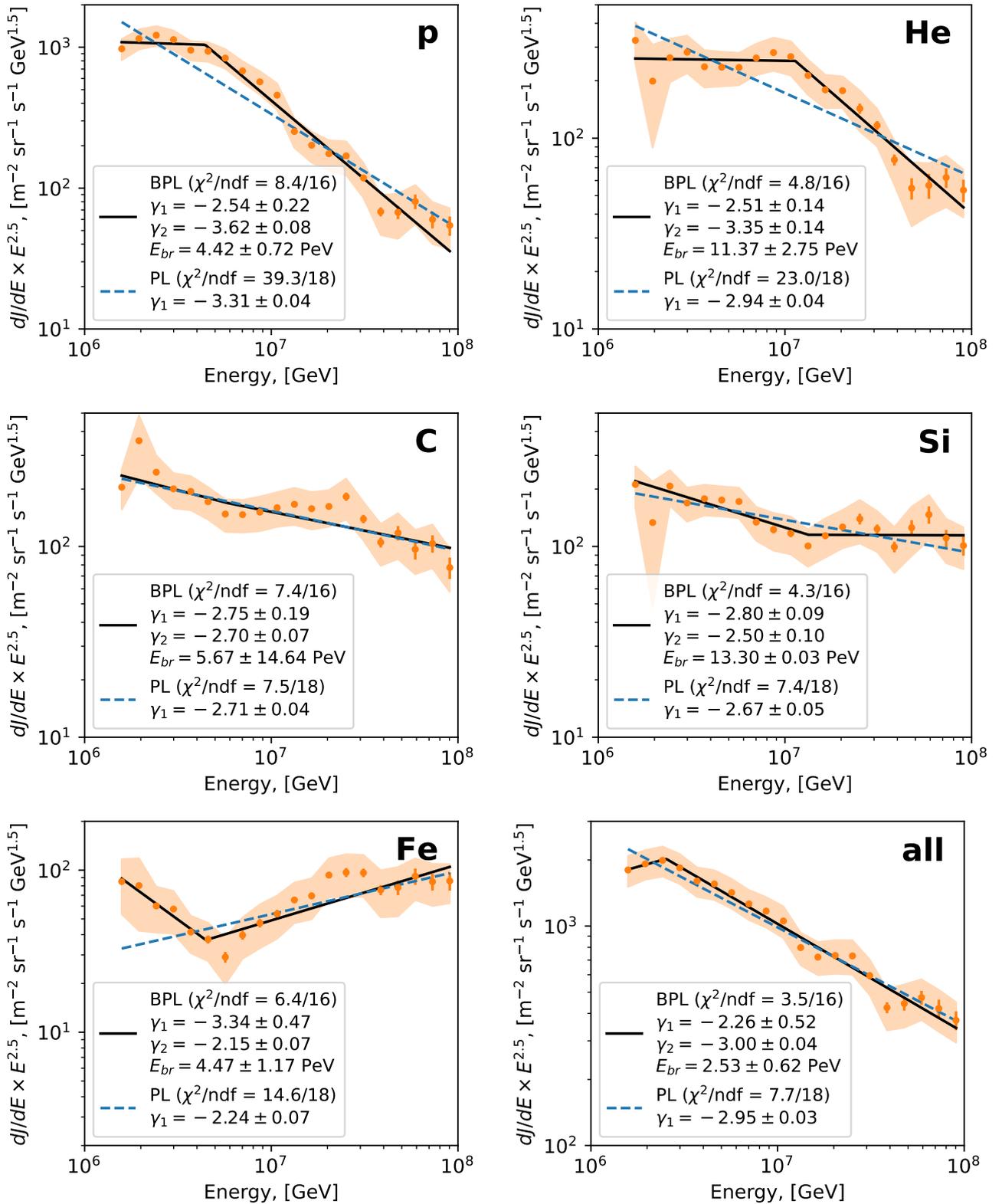


Рисунок 3.8 — Аппроксимации энергетических спектров (слева направо сверху вниз): протонной, гелиевой, углеродной, кремниевой, железной массовых компонент и полного спектра простой (PL, синий пунктир) и ломаной (BPL, чёрная линия) степенными функциями

при $\sim 4,4 \text{ ПэВ} \times 14 = \sim 60 \text{ ПэВ}$, но выше этой энергии в данном исследовании доступны лишь две энергетические точки, причём с малой статистикой и большей систематической неопределённостью относительно низких энергий и поэтому недоступно для значимого наблюдения; а в случае углерода при $\sim 4,4 \text{ ПэВ} \times 6 = \sim 25 \text{ ПэВ}$, и в целом, при внимательном рассмотрении углеродного спектра на Рисунке 3.8 с учётом только статистических неопределённостей в этом месте действительно наблюдается намёк на коленоподобное поведение, его наблюдение на значимом уровне могло бы стать важным доказательством в зарядовую зависимость положения колена, в том числе поэтому дальнейшая работа по уменьшению систематических неопределённостей спектров индивидуальных компонент оправдана.

Наконец, особый интерес представляет впервые обнаруженное в данной работе указание на «ужесточение» в энергетическом спектре железной компоненты при энергии $\sim 4,5 \text{ ПэВ}$. Примечательно, что его положение соответствует «ужесточению», обнаруженному ранее в эксперименте GRAPES-3 в спектре протонов при энергии 166 ТэВ [88], в предположении его зависимости от заряда первичной частицы ($166 \text{ ТэВ} \times 26 = 4,4 \text{ ПэВ}$). Однако поскольку в данной работе получено лишь указание на «ужесточение» в железном спектре (со значимостью $2,4 \sigma$), то необходимы дальнейшие эксперименты для уточнения результатов, причём как для спектров железной, так и других массовых компонент (в соответствующих энергетических диапазонах), поскольку в них должны проявляться те же свойства. Наиболее перспективный результат может предоставить эксперимент LHAASO, энергетический диапазон которого покрывает предполагаемые энергетические положения «ужесточения» для различных массовых компонент, а систематические неопределённости ниже чем в других современных экспериментах.

Заключение

Основные результаты работы заключаются в следующем.

1. Разработана методика восстановления энергетических спектров пяти индивидуальных массовых компонент (протоны, гелий, углерод, кремний, железо) в диапазоне энергий 1,5–100 ПэВ на архивных данных эксперимента KASCADE в диапазоне энергий 1,5–100 ПэВ с использованием классификатора на основе машинного обучения для пособытийной реконструкции массовой компоненты первичной частицы, а также последовательного анфолдинга реконструированных массовой компоненты и энергии первичной частицы. При прямом сравнении продемонстрировано, что точность восстановления массового состава методикой, разработанной в данной работе, выше, чем в оригинальном исследовании KASCADE, а также в современных экспериментах (IceTop).
2. Получены энергетические спектры пяти массовых компонент в диапазоне энергий от 1,5 ПэВ до 100 ПэВ с использованием Монте-Карло моделирования, основанном на современных генераторах адронных взаимодействий QGSJet-II.04, EPOS-LHC, Sibyll 2.3c. Дополнительно результат продемонстрирован в виде среднего логарифма массового числа в зависимости от энергии первичной частицы. Отдельно учтены «теоретические» неопределённости, связанные с выбором генератора.
3. Продемонстрировано доминирующее поведение спектра протонной компоненты над остальными в диапазоне энергий 1,5–10 ПэВ (~ 10 раз при энергии 10 ПэВ), находящееся в противоречии с оригинальными результатами KASCADE, в которых доминирует гелиевая компонента.
4. Проведено сравнение полученных в данной работе результатов с последними результатами экспериментов IceTop, TALE и LHAASO. Показано согласие результатов с ними в рамках «теоретических» неопределённостей.
5. Исследовано поведение энергетических спектров индивидуальных массовых компонент. Впервые обнаружено статистически значимое колленоподобное поведение в случае протонной (более 5σ) и гелиевой (более 3σ) компонент при энергиях $\sim 4,4$ ПэВ и ~ 11 ПэВ соответственно.

Впервые получено указание ($2,4 \sigma$) на излом в железной компоненте при энергии $\sim 4,5$ ПэВ.

Автор выражает благодарность своему научному руководителю Кузнецову М. Ю. за поддержку и помощь на всех этапах исследования и написания работы. Также автор благодарит Костюнина Д. Г. за идею исследования и всестороннюю помощь в процессе работы; выражает признательность Сотникову В. В., Леноку В. В., Плохих И. А. за активный и неформальный процесс обсуждения результатов; благодарит команду KCDC за качественные экспериментальные данные и Монте-Карло моделирование KASCADE, выложенные в открытый доступ, исчерпывающую документацию к ним, а также за дополнительно сгенерированное Монте-Карло для целей текущей работы; благодарит коллег из Лаборатории 2 Института ядерной физики им. Г. И. Будкера, в частности Лукина П. А., за переданные знания, оказавшиеся востребованными в текущем исследовании.

Список сокращений и условных обозначений

ATLAS	A Toroidal LHC Apparatus (многоцелевой эксперимент по физике элементарных частиц на Большом адронном коллайдере)
CNN	Convolutional Neural Network (свёрточная нейронная сеть)
CORSIKA	COsmic Ray SIMulations for KAScade (программа для детального моделирования ШАЛов, инициированных первичными космическими лучами)
CRES	Cosmic Ray Event Simulation (программный пакет для моделирования сигналов и энерговыделений в детекторах KASCADE)
EfficientNet v2	семейство оптимизированных свёрточных нейронных сетей
EPOS	Energy conserving quantum mechanical approach, based on Partons, parton ladders, strings, Off-shell remnants, and Splitting of parton ladders (семейство генераторов адронных взаимодействий высоких энергий)
FLUKA	FLUktuierende KAskade (пакет программ для моделирования прохождения частиц через вещество методами Монте-Карло)
GEANT 3	GEometry ANd Tracking (программный пакет для моделирования прохождения элементарных частиц через вещество с использованием методов Монте-Карло)
GRAPES-3	Gamma Ray Astronomy PeV EnergieS phase-3 (эксперимент по изучению космических лучей в Утакаманде, Индия)
KASCADE	KARlsruhe Shower Core and Array DETector (эксперимент по изучению космических лучей в Карлсруэ, Германия)
KCDC	KASCADE Cosmic Ray Data Centre (веб-портал с опубликованными данными эксперимента KASCADE)
KRETA	KASCADE Reconstruction for ExTensiveAir showers (программа реконструкции основных параметров ШАЛа в эксперименте KASCADE)
LHAASO	Large High Altitude Air Shower Observatory (обсерватория по изучению гамма-излучения и космических лучей в Даочэне, Китай)
MLP	MultiLayer Perceptron (многослойный перцептрон)

QGSJet	Quark-Gluon String model with Jets (семейство генераторов адронных взаимодействий высоких энергий)
RF	Random Forest classifier (классификатор методом случайного леса)
TAIGA	Tunka Advanced Instrument for cosmic rays and Gamma Astronomy (гибридный эксперимент для измерения ШАЛов в диапазоне энергий от ТэВ до ПэВ)
TALE	Telescope Array Low-energy Extension (низкоэнергетическое расширение эксперимента Telescope Array)
TOTEM	TOTAL Elastic and diffractive cross section Measurement (эксперимент по изучению взаимодействий протонов под малыми углами на Большом адронном коллайдере)
ФЭУ	ФотоЭлектронный Умножитель
ШАЛ	Широкий Атмосферный Ливень

Список литературы

1. *Hess, V. F.* Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten / V. F. Hess. — Текст: электронный // *Phys. Z.* — 1912. — Т. 13. — С. 1084—1091. — URL: <https://www.mpi-hd.mpg.de/hfm/HESS/public/HessArticle.pdf> (дата обращения: 28.05.2024).
2. The origin of Galactic cosmic rays: Challenges to the standard paradigm / S. Gabici [и др.]. — Текст: электронный // *International Journal of Modern Physics D.* — 2019. — Т. 28, № 15. — С. 1930022. — URL: <https://doi.org/10.1142/S0218271819300222> (дата обращения: 29.05.2024).
3. *Becker Tjus, J.* Closing in on the origin of Galactic cosmic rays using multimessenger information / J. Becker Tjus [и др.]. — Текст: электронный // *Physics Reports.* — 2020. — Т. 872. — С. 1—98. — URL: <https://www.sciencedirect.com/science/article/pii/S0370157320301927> (дата обращения: 29.05.2024).
4. *Blasi, P.* The Origin of Galactic Cosmic Rays / P. Blasi. — Текст: электронный // *Astron. Astrophys. Rev.* — 2013. — Т. 21. — С. 70. — URL: <https://doi.org/10.1007/s00159-013-0070-7> (дата обращения: 30.05.2024).
5. *Amato, E.* The origin of galactic cosmic rays / E. Amato. — Текст: электронный // *International Journal of Modern Physics D.* — 2014. — Т. 23, № 07. — С. 1430013. — URL: <https://doi.org/10.1142/S0218271814300134> (дата обращения: 30.05.2024).
6. *Sciascio, G. D.* Measurement of Energy Spectrum and Elemental Composition of PeV Cosmic Rays: Open Problems and Prospects / G. D. Sciascio. — Текст: электронный // *Appl. Sciences.* — 2022. — Т. 12, № 2. — С. 705. — URL: <https://doi.org/10.3390/app12020705> (дата обращения: 13.08.2024).
7. *Kulikov, G. V.* On the Size Spectrum of Extensive Air Showers / G. V. Kulikov [и др.]. — Текст: электронный // *Journal of Experimental and Theoretical Physics.* — 1959. — Т. 8, № 3. — С. 441. — URL: http://jetp.ras.ru/cgi-bin/dn/e_008_03_0441.pdf (дата обращения: 29.05.2024).

8. The Cosmic ray energy spectrum between $10^{14.5}$ -eV and $10^{16.3}$ -eV covering the 'knee' region / М. Amenomori [и др.]. — Текст: электронный // *Astrophys. J.* — 1995. — Т. 461. — С. 408—414. — URL: <https://doi.org/10.1086/177069> (дата обращения: 29.05.2024).
9. Energy spectrum and chemical composition of cosmic rays between 0.3-PeV and 10-PeV determined from the Cherenkov light and charged particle distributions in air showers / F. Arqueros [и др.]. — Текст: электронный // *Astron. Astrophys.* — 2000. — Т. 359. — С. 682—694. — URL: <https://ui.adsabs.harvard.edu/abs/2000A&A...359..682H> (дата обращения: 29.05.2024).
10. Energy spectrum of primary cosmic rays between $10^{14.5}$ and 10^{18} eV / М. Nagano [и др.]. — Текст: электронный // *Journal of Physics G: Nuclear Physics.* — 1984. — Сент. — Т. 10, № 9. — С. 1295. — URL: <https://dx.doi.org/10.1088/0305-4616/10/9/016> (дата обращения: 29.05.2024).
11. The cosmic ray energy spectrum between 10^{14} and 10^{16} eV / М. Glasmacher [и др.]. — Текст: электронный // *Astroparticle Physics.* — 1999. — Т. 10, № 4. — С. 291—302. — URL: <https://www.sciencedirect.com/science/article/pii/S092765059800070X> (дата обращения: 29.05.2024).
12. *Swordy, S. P.* Elemental composition of cosmic rays near the knee by multiparameter measurements of air showers / S. P. Swordy [и др.]. — Текст: электронный // *Astroparticle Physics.* — 2000. — Т. 13, № 2. — С. 137—150. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650599001176> (дата обращения: 29.05.2024).
13. KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems / Т. Antoni [и др.]. — Текст: электронный // *Astroparticle Physics.* — 2005. — Т. 24, № 1. — С. 1—25. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650505000691> ; Дата публикации: 08.06.2005.
14. Cosmic ray composition and energy spectrum from 1 – 30 PeV using the 40-string configuration of IceTop and IceCube / R. Abbasi [и др.]. — Текст: электронный // *Astroparticle Physics.* — 2013. — Т. 42. — С. 15—32. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650512002009> (дата обращения: 29.05.2024).

15. *Hörandel, J. R.* Models of the knee in the energy spectrum of cosmic rays / J. R. Hörandel. — Текст: электронный // *Astroparticle Physics*. — 2004. — Т. 21, № 3. — С. 241—265. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650504000209> (дата обращения: 30.05.2024).
16. *Kachelrieß, M.* Cosmic ray models / М. Kachelrieß [и др.]. — Текст: электронный // *Progress in Particle and Nuclear Physics*. — 2019. — Т. 109. — С. 103710. — URL: <https://doi.org/10.1016/j.pnpnp.2019.07.002> (дата обращения: 15.06.2024).
17. *Peters, B.* Primary cosmic radiation and extensive air showers / B. Peters. — Текст: электронный // *Nuovo Cim.* — 1961. — Т. 22, № 4. — С. 800—819. — URL: <https://doi.org/10.1007/bf02783106> (дата обращения: 30.05.2024).
18. *Hillas, A.* Where do 10^{19} eV cosmic rays come from? / A. Hillas. — Текст: электронный // *Nuclear Physics B - Proceedings Supplements*. — 2004. — Т. 136. — С. 139—146. — URL: <https://www.sciencedirect.com/science/article/pii/S092056320400461X> (дата обращения: 30.05.2024) ; CRIS 2004 Proceedings of the Cosmic Ray International Seminars: GZK and Surroundings.
19. *Blasi, P.* Diffusive propagation of cosmic rays from supernova remnants in the Galaxy. I: spectrum and chemical composition / P. Blasi [и др.]. — Текст: электронный // *JCAP*. — 2012. — Т. 01. — С. 010. — URL: <https://doi.org/10.1088/1475-7516/2012/01/010> (дата обращения: 30.05.2024).
20. *Giacinti, G.* Escape model for Galactic cosmic rays and an early extragalactic transition / G. Giacinti [и др.]. — Текст: электронный // *Phys. Rev. D*. — 2015. — Апр. — Т. 91, вып. 8. — С. 083009. — URL: <https://link.aps.org/doi/10.1103/PhysRevD.91.083009> (дата обращения: 30.05.2024).
21. *Dar, A.* A theory of cosmic rays / A. Dar [и др.]. — Текст: электронный // *Physics Reports*. — 2008. — Т. 466, № 6. — С. 179—241. — URL: <https://doi.org/10.1016/j.physrep.2008.05.004> (дата обращения: 15.06.2024).
22. Are protons still dominant at the knee of the cosmic-ray energy spectrum? / М. Amenomori [и др.]. — Текст: электронный // *Phys. Lett. B*. — 2006. — Т. 632. — С. 58—64. — URL: <https://doi.org/10.1016/j.physletb.2005.10.048> (дата обращения: 30.05.2024).

23. KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays / W. D. Apel [и др.]. — Текст: электронный // *Astroparticle Physics*. — 2013. — Т. 47. — С. 54—66. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650513000996> ; Дата публикации: 22.06.2013.
24. Cosmic ray spectrum and composition from PeV to EeV using 3 years of data from IceTop and IceCube / M. G. Aartsen [и др.]. — Текст: электронный // *Phys. Rev. D*. — 2019. — Окт. — Т. 100, вып. 8. — С. 082002. — URL: <https://link.aps.org/doi/10.1103/PhysRevD.100.082002> (дата обращения: 29.03.2024).
25. The Cosmic-Ray Composition between 2 PeV and 2 EeV Observed with the TALE Detector in Monocular Mode / R. U. Abbasi [и др.]. — Текст: электронный // *The Astrophysical Journal*. — 2021. — Март. — Т. 909, № 2. — С. 178. — URL: <https://dx.doi.org/10.3847/1538-4357/abdd30> (дата обращения: 29.03.2024).
26. Study of extensive air showers and primary energy spectra by МАКЕТ-ANI detector on mountain Aragats / A. Chilingarian [и др.]. — Текст: электронный // *Astroparticle Physics*. — 2007. — Т. 28, № 1. — С. 58—71. — URL: <https://www.sciencedirect.com/science/article/pii/S0927650507000552> (дата обращения: 23.07.2024).
27. Tunka-133: Results of 3 year operation / V. V. Prosin [и др.]. — Текст: электронный // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. — 2014. — Т. 756. — С. 94—101. — URL: <https://www.sciencedirect.com/science/article/pii/S0168900213012527> (дата обращения: 30.05.2024).
28. Cosmic-Ray Research at the TAIGA Astrophysical Facility: Results and Plans / I. I. Astarov [и др.]. — Текст: электронный // *J. Exp. Theor. Phys.* — 2022. — Т. 134, № 4. — С. 469—478. — URL: <https://doi.org/10.1134/S1063776122040136> (дата обращения: 30.05.2024).
29. Measurements of All-Particle Energy Spectrum and Mean Logarithmic Mass of Cosmic Rays from 0.3 to 30 PeV with LHAASO-KM2A / Z. Cao [и др.]. — Текст: электронный // *Phys. Rev. Lett.* — 2024. — Март. — Т. 132, вып. 13. — С. 131002. — URL: <https://link.aps.org/doi/10.1103/PhysRevLett.132.131002> ; Дата публикации: 26.03.2024.

30. *Tomassetti, N.* Direct Measurements of Galactic Cosmic Rays / N. Tomassetti. — Текст: электронный // PoS. — 2023. — Т. ECRS. — С. 007. — URL: <https://doi.org/10.22323/1.423.0007> (дата обращения: 30.05.2024).
31. *Engel, R.* Indirect Detection of Cosmic Rays / R. Engel. — Текст: электронный // Handbook of Particle Detection and Imaging / под ред. С. Grupen [и др.]. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2012. — С. 593—632. — URL: https://doi.org/10.1007/978-3-642-13271-1_24 (дата обращения: 30.05.2024).
32. Extensive Cosmic-Ray Showers / P. Auger [и др.]. — Текст: электронный // Rev. Mod. Phys. — 1939. — Июль. — Т. 11, вып. 3/4. — С. 288—291. — URL: <https://link.aps.org/doi/10.1103/RevModPhys.11.288> (дата обращения: 30.05.2024).
33. The Cosmic Ray Energy Spectrum between 2 PeV and 2 EeV Observed with the TALE Detector in Monocular Mode / R. U. Abbasi [и др.]. — Текст: электронный // The Astrophysical Journal. — 2018. — Сент. — Т. 865, № 1. — С. 74. — URL: <https://dx.doi.org/10.3847/1538-4357/aada05> (дата обращения: 30.05.2024).
34. IceTop: The surface component of IceCube / R. Abbasi [и др.]. — Текст: электронный // Nucl. Instrum. Meth. A. — 2013. — Т. 700. — С. 188—220. — URL: <https://doi.org/10.1016/j.nima.2012.10.067> (дата обращения: 30.05.2024).
35. *Cao, Z.* A future project at Tibet: The large high altitude air shower observatory (LHAASO) / Z. Cao. — Текст: электронный // Chin. Phys. C. — 2010. — Т. 34. — С. 249—252. — URL: <https://doi.org/10.1088/1674-1137/34/2/018> (дата обращения: 30.05.2024).
36. Modernization of the Carpet-2 array of the Baksan Neutrino Observatory / D. D. Dzharruev [и др.]. — Текст: электронный // Bull. Russ. Acad. Sci. Phys. — 2007. — Т. 71, № 4. — С. 525—527. — URL: <https://doi.org/10.3103/S1062873807040260> (дата обращения: 08.06.2024).
37. *Pierog, T.* Hadronic Interactions and Air Showers: Where Do We Stand? / T. Pierog. — Текст: электронный // EPJ Web Conf. / под ред. В. Pattison [и др.]. — 2019. — Т. 208. — С. 02002. — URL: <https://doi.org/10.1051/epjconf/201920802002> (дата обращения: 30.05.2024).

38. *Ostapchenko, S.* Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model / S. Ostapchenko. — Текст: электронный // Phys. Rev. D. — 2011. — Т. 83. — С. 014018. — URL: <https://doi.org/10.1103/PhysRevD.83.014018> (дата обращения: 30.05.2024).
39. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider / T. Pierog [и др.]. — Текст: электронный // Phys. Rev. C. — 2015. — Сент. — Т. 92, вып. 3. — С. 034906. — URL: <https://link.aps.org/doi/10.1103/PhysRevC.92.034906> ; Дата публикации: 14.09.2015.
40. SIBYLL: An Event generator for simulation of high-energy cosmic ray cascades / R. S. Fletcher [и др.]. — Текст: электронный // Phys. Rev. D. — 1994. — Т. 50. — С. 5710—5731. — URL: <https://doi.org/10.1103/PhysRevD.50.5710> (дата обращения: 30.05.2024).
41. The cosmic-ray experiment KASCADE / T. Antoni [и др.]. — Текст: электронный // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. — 2003. — Т. 513, № 3. — С. 490—510. — URL: [https://doi.org/10.1016/S0168-9002\(03\)02076-X](https://doi.org/10.1016/S0168-9002(03)02076-X) ; Дата публикации: 11.11.2003.
42. The KASCADE Cosmic-ray Data Centre KCDC: Granting Open Access to Astroparticle Physics Research Data / A. Haungs [и др.]. — Текст: электронный // Eur. Phys. J. C. — 2018. — Т. 78, № 9. — С. 741. — URL: <https://doi.org/10.1140/epjc/s10052-018-6221-2> ; Дата публикации: 17.09.2018.
43. Using deep learning to enhance event geometry reconstruction for the telescope array surface detector / D. Ivanov [и др.]. — Текст: электронный // Mach. Learn. Sci. Tech. — 2021. — Т. 2, № 1. — С. 015006. — URL: <https://doi.org/10.1088/2632-2153/abae74> (дата обращения: 31.05.2024).
44. New insights from old cosmic rays: A novel analysis of archival KASCADE data / D. Kostunin [и др.]. — Текст: электронный // PoS. — 2021. — Т. ICRC2021. — С. 319. — URL: <https://doi.org/10.22323/1.395.0319> (дата обращения: 31.05.2024).
45. *Erdmann, M.* A deep learning-based reconstruction of cosmic ray-induced air showers / M. Erdmann [и др.]. — Текст: электронный // Astropart. Phys. — 2018. — Т. 97. — С. 46—53. — URL: <https://doi.org/10.1016/j.astropartphys.2017.10.006> (дата обращения: 31.05.2024).

46. Mass composition of ultrahigh-energy cosmic rays with the Telescope Array Surface Detector data / R. U. Abbasi [и др.]. — Текст: электронный // Phys. Rev. D. — 2019. — Т. 99, № 2. — С. 022002. — URL: <https://doi.org/10.1103/PhysRevD.99.022002> (дата обращения: 31.05.2024).
47. Deep-learning based reconstruction of the shower maximum X_{max} using the water-Cherenkov detectors of the Pierre Auger Observatory / A. Aab [и др.]. — Текст: электронный // JINST. — 2021. — Т. 16, № 07. — P07019. — URL: <https://doi.org/10.1088/1748-0221/16/07/P07019> (дата обращения: 31.05.2024).
48. Deep learning method for identifying mass composition of ultra-high-energy cosmic rays / O. Kalashev [и др.]. — Текст: электронный // JINST. — 2022. — Т. 17, № 05. — P05008. — URL: <https://doi.org/10.1088/1748-0221/17/05/P05008> (дата обращения: 31.05.2024).
49. Method for Separating Extensive Air Showers by Primary Mass Using Machine Learning for a Sphere-Type Cherenkov Telescope / V. S. Latypova [и др.]. — Текст: электронный // Moscow Univ. Phys. Bull. — 2023. — Т. 78, Suppl 1. — S25—S31. — URL: <https://doi.org/10.3103/S0027134923070196> (дата обращения: 31.05.2024).
50. Cosmic-Ray Composition with IceTop and IceCube using graph neural networks / P. Koundal [и др.]. — Текст: электронный // PoS. — 2023. — Т. ICRC2023. — С. 334. — URL: <https://doi.org/10.22323/1.444.0334> (дата обращения: 31.05.2024).
51. Jin, C. Classifying cosmic-ray proton and light groups in LHAASO-KM2A experiment with graph neural network / C. Jin [и др.]. — Текст: электронный // Chin. Phys. C. — 2020. — Т. 44, № 6. — С. 065002. — URL: <https://doi.org/10.1088/1674-1137/44/6/065002> (дата обращения: 31.05.2024).
52. Methods of machine learning for the analysis of cosmic rays mass composition with the KASCADE experiment data / M. Y. Kuznetsov, N. A. Petrov, I. A. Plokhikh, V. V. Sotnikov. — Текст: электронный // Journal of Instrumentation. — 2024. — Янв. — Т. 19, № 01. — P01025. — URL: <https://doi.org/10.1088/1748-0221/19/01/P01025> ; Дата публикации: 25.01.2024.

53. Energy spectra of elemental groups of cosmic rays with the KASCADE experiment data and machine learning / M. Yu. Kuznetsov, N. A. Petrov, I. A. Plokhikh, V. V. Sotnikov. — Текст: электронный // Journal of Cosmology and Astroparticle Physics. — 2024. — Май. — Т. 2024, № 05. — С. 125. — URL: <https://doi.org/10.1088/1475-7516/2024/05/125> ; Дата публикации: 30.05.2024.
54. Towards mass composition study with KASCADE using deep neural networks / V. Sotnikov, M. Kuznetsov, N. Petrov, I. Plokhikh. — Текст: электронный // Proceedings of 27th European Cosmic Ray Symposium — PoS(ECRS). Т. 423. — 2023. — С. 092. — URL: <https://doi.org/10.22323/1.423.0092> ; Дата публикации: 14.12.2023.
55. The KASCADE-Grande experiment / W. D. Apel [и др.]. — Текст: электронный // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. — 2010. — Т. 620, № 2. — С. 202—216. — URL: <https://doi.org/10.1016/j.nima.2010.03.147> ; Дата публикации: 08.04.2010.
56. KCDC User Manual: Open Access Solution for the KASCADE / J. Wochele [и др.]. — Текст: электронный. — re3data.org - Registry of Research Data Repositories, 11.2013. — URL: <https://doi.org/10.17616/R3TS4P> (дата обращения: 15.02.2024).
57. *Ostapchenko, S.* Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model / S. Ostapchenko. — Текст: электронный // Phys. Rev. D. — 2011. — Янв. — Т. 83, вып. 1. — С. 014018. — URL: <https://link.aps.org/doi/10.1103/PhysRevD.83.014018> ; Дата публикации: 25.01.2011.
58. The hadronic interaction model Sibyll 2.3c and Feynman scaling / F. Riehn [и др.]. — Текст: электронный // Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017). Т. 301. — 2017. — С. 301. — URL: <https://doi.org/10.22323/1.301.0301> ; Дата публикации: 03.08.2018.
59. FLUKA: A multi-particle transport code (program version 2005) / A. Ferrari [и др.]. — Текст: электронный. — Geneva : CERN, 2005. — (CERN Yellow Reports: Monographs). — URL: <http://cds.cern.ch/record/898301> ; Дата публикации: 12.10.2005.

60. GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version / R. Brun [и др.]. — Текст: электронный. — Geneva : CERN, 1987. — URL: <http://cds.cern.ch/record/1119728> ; Дата публикации: 09.09.1987.
61. *Ostapchenko, S.* QGSJET-II: towards reliable description of very high energy hadronic interactions / S. Ostapchenko. — Текст: электронный // Nuclear Physics B - Proceedings Supplements. — 2006. — Т. 151, № 1. — С. 143—146. — URL: <https://www.sciencedirect.com/science/article/pii/S0920563205009175> ; Дата публикации: 10.01.2006.
62. *Finger, M. R.* Reconstruction of energy spectra for different mass groups of high-energy cosmic rays : дис. ... канд. / Finger Marcel René. — 2011. — URL: <https://doi.org/10.5445/IR/1000023830> ; Дата публикации: 27.05.2011.
63. *Lenok, V.* Measurement of the Cosmic-Ray Energy Spectrum Using a Novel Approach to Model the Aperture of Radio Arrays : дис. ... канд. / Lenok Vladimir. — Karlsruher Institut für Technologie (KIT), 2022. — 147 с. — URL: <https://doi.org/10.5445/IR/1000143479> ; Дата публикации: 16.03.2022.
64. *Cowan, G.* Statistical Data Analysis / G. Cowan. — Текст: электронный. — Clarendon Press, 1998. — (Oxford science publications). — URL: <https://books.google.ru/books?id=ff8ZyW0nlJAC> ; Дата публикации: 1998.
65. *Ho, T. K.* Random decision forests / Т. К. Но. — Текст: электронный // Proceedings of 3rd International Conference on Document Analysis and Recognition. Т. 1. — 1995. — 278—282 vol.1. — URL: <https://doi.org/10.1109/ICDAR.1995.598994> ; Дата публикации: 16.08.1995.
66. *Breiman, L.* Random Forests / L. Breiman. — Текст: электронный // Machine Learning. — 2001. — Т. 45, № 1. — С. 5—32. — URL: <http://dx.doi.org/10.1023/A:1010933404324> ; Дата публикации: 2001.
67. *Tan, M.* Efficientnetv2: Smaller models and faster training / М. Тан [и др.]. — Текст: электронный // International conference on machine learning. — PMLR. 2021. — С. 10096—10106. — URL: <https://api.semanticscholar.org/CorpusID:232478903> (дата обращения: 27.02.2024).

68. Scikit-learn: Machine Learning in Python / F. Pedregosa [и др.]. — Текст: электронный // Journal of Machine Learning Research. — 2011. — Т. 12. — С. 2825—2830. — URL: <https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html> (дата обращения: 27.02.2024).
69. *Ioffe, S.* Batch normalization: Accelerating deep network training by reducing internal covariate shift / S. Ioffe [и др.]. — Текст: электронный // Proceedings of the 32nd International Conference on International Conference on Machine Learning. Т. 37. — PMLR. Lille, France : JMLR.org, 2015. — С. 448—456. — (ICML'15). — URL: <http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43442.pdf> (дата обращения: 27.02.2024).
70. *Clevert, D.* Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs) / D. Clevert [и др.]. — Текст: электронный // 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings / под ред. Y. Bengio [и др.]. — 2016. — URL: <http://arxiv.org/abs/1511.07289> (дата обращения: 28.02.2024).
71. Dropout: A Simple Way to Prevent Neural Networks from Overfitting / N. Srivastava [и др.]. — Текст: электронный // Journal of Machine Learning Research. — 2014. — Т. 15, № 56. — С. 1929—1958. — URL: <http://jmlr.org/papers/v15/srivastava14a.html> (дата обращения: 28.02.2024).
72. *Kingma, D. P.* Adam: A Method for Stochastic Optimization / D. P. Kingma [и др.]. — Текст: электронный // 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. — 2015. — URL: <https://arxiv.org/abs/1412.6980> (дата обращения: 27.02.2024).
73. TensorFlow: A System for Large-Scale Machine Learning / М. Abadi [и др.]. — Текст: электронный // 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16). — Savannah, GA : USENIX Association, 11.2016. — С. 265—283. — URL: <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi> (дата обращения: 27.02.2024).
74. Gradient-based learning applied to document recognition / Y. Lecun [и др.]. — Текст: электронный // Proceedings of the IEEE. — 1998. — Т. 86, № 11. — С. 2278—2324. — URL: <https://doi.org/10.1109/5.726791> ; Дата публикации: 17.04.1998.

75. *Glorot, X.* Deep Sparse Rectifier Neural Networks / X. Glorot [и др.]. — Текст: электронный // Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. Т. 15 / под ред. G. Gordon [и др.]. — Fort Lauderdale, FL, USA : PMLR, 04.2011. — С. 315—323. — (Proceedings of Machine Learning Research). — URL: <https://proceedings.mlr.press/v15/glorot11a.html> (дата обращения: 28.02.2024).
76. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke [и др.]. — Текст: электронный // Advances in Neural Information Processing Systems. Т. 32 / под ред. H. Wallach [и др.]. — Curran Associates, Inc., 2019. — С. 8026—8037. — URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf (дата обращения: 28.02.2024).
77. *Biewald, L.* Experiment Tracking with Weights and Biases / L. Biewald. — Электронный ресурс. — 2020. — URL: <https://www.wandb.com/> (дата обращения: 28.02.2024) ; Software available from wandb.com.
78. *Matthews, J.* A Heitler model of extensive air showers / J. Matthews. — Текст: электронный // Astropart. Phys. — 2005. — Т. 22. — С. 387—397. — URL: <https://doi.org/10.1016/j.astropartphys.2004.09.003> (дата обращения: 10.08.2024).
79. *Letessier-Selvon, A.* Ultrahigh Energy Cosmic Rays / A. Letessier-Selvon [и др.]. — Текст: электронный // Rev. Mod. Phys. — 2011. — Т. 83. — С. 907—942. — URL: <https://doi.org/10.1103/RevModPhys.83.907> (дата обращения: 13.08.2024).
80. First measurement of elastic, inelastic and total cross-section at $\sqrt{s} = 13$ TeV by TOTEM and overview of cross-section data at LHC energies / G. Antchev [и др.]. — Текст: электронный // Eur. Phys. J. C. — 2019. — Т. 79, № 2. — С. 103. — URL: <https://doi.org/10.1140/epjc/s10052-019-6567-0> (дата обращения: 05.08.2024).
81. Measurement of the total cross section from elastic scattering in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector / M. Aaboud [и др.]. — Текст: электронный // Phys. Lett. B. — 2016. — Т. 761. — С. 158—178. — URL: <https://doi.org/10.1016/j.physletb.2016.08.020> (дата обращения: 05.08.2024).

82. *Ostapchenko, S.* High energy interactions of cosmic rays / S. Ostapchenko. — Текст: электронный // *Advances in Space Research*. — 2019. — Т. 64, № 12. — С. 2445—2458. — URL: <https://doi.org/10.1016/j.asr.2019.05.050> (дата обращения: 30.08.2024) ; *Advances in Cosmic-Ray Astrophysics and Related Areas*.
83. The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider / J. Albrecht [и др.]. — Текст: электронный // *Astrophys. Space Sci.* — 2022. — Т. 367, № 3. — С. 27. — URL: <https://doi.org/10.1007/s10509-022-04054-5> (дата обращения: 15.08.2024).
84. *D'Agostini, G.* A multidimensional unfolding method based on Bayes' theorem / G. D'Agostini. — Текст: электронный // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. — 1995. — Т. 362, № 2. — С. 487—498. — URL: [https://doi.org/10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X) (дата обращения: 27.03.2024).
85. *Bourbeau, J.* PyUnfold: A Python package for iterative unfolding / J. Bourbeau [и др.]. — Текст: электронный // *The Journal of Open Source Software*. — 2018. — Июнь. — Т. 3, № 26. — С. 741. — URL: <https://doi.org/10.21105/joss.00741> (дата обращения: 27.03.2024).
86. Comparison of unfolding methods using RooFitUnfold / L. Brenner [и др.]. — Текст: электронный // *Int. J. Mod. Phys. A*. — 2020. — Т. 35, № 24. — С. 2050145. — URL: <https://doi.org/10.1142/S0217751X20501456> ; Дата публикации: 31.08.2020.
87. *James, F.* Minuit - a system for function minimization and analysis of the parameter errors and correlations / F. James [и др.]. — Текст: электронный // *Computer Physics Communications*. — 1975. — Т. 10, № 6. — С. 343—367. — URL: <https://www.sciencedirect.com/science/article/pii/0010465575900399> (дата обращения: 22.04.2024).
88. Evidence of a Hardening in the Cosmic Ray Proton Spectrum at around 166 TeV Observed by the GRAPES-3 Experiment / F. Varsi [и др.]. — Текст: электронный // *Phys. Rev. Lett.* — 2024. — Янв. — Т. 132, вып. 5. — С. 051002. — URL: <https://doi.org/10.1103/PhysRevLett.132.051002> ; Дата публикации: 31.01.2024.

Приложение А

Численные значения результатов

В этом приложении представлены таблицы с полученными в данной работе результатами. В Таблице 4 показан полный дифференциальный поток в зависимости от энергии первичной частицы; в Таблицах 5 — 9 — дифференциальные потоки для пяти индивидуальных массовых компонент (p , He , C , Si , Fe) в зависимости от энергии первичной частицы. Наконец, в Таблице 10 представлены результаты зависимости $\langle \ln A \rangle$ от энергии первичной частицы.

Таблица 4 — Полученные в данном исследовании значения дифференциального полного потока (dJ_{all}/dE , [$\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}$]) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$dJ_{all}/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ [$\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}$]
6,15–6,2425	$\left(5,738 \pm 0,012 \pm 0,931 / {}^{+1,073}_{-1,384}\right) \times 10^{-13}$
6,2425–6,335	$\left(3,594 \pm 0,009 \pm 0,549 / {}^{+0,869}_{-0,734}\right) \times 10^{-13}$
6,335–6,4275	$\left(2,189 \pm 0,006 \pm 0,342 / {}^{+0,404}_{-0,555}\right) \times 10^{-13}$
6,4275–6,52	$\left(1,188 \pm 0,004 \pm 0,185 / {}^{+0,185}_{-0,369}\right) \times 10^{-13}$
6,52–6,6125	$\left(6,073 \pm 0,029 \pm 0,929 / {}^{+1,287}_{-1,476}\right) \times 10^{-14}$
6,6125–6,705	$\left(3,459 \pm 0,019 \pm 0,535 / {}^{+0,535}_{-1,239}\right) \times 10^{-14}$
6,705–6,7975	$\left(1,855 \pm 0,013 \pm 0,286 / {}^{+0,307}_{-0,550}\right) \times 10^{-14}$
6,7975–6,89	$\left(9,678 \pm 0,085 \pm 1,510 / {}^{+1,510}_{-2,922}\right) \times 10^{-15}$
6,89–6,9825	$\left(5,267 \pm 0,057 \pm 0,834 / {}^{+0,834}_{-1,512}\right) \times 10^{-15}$
6,9825–7,075	$\left(2,787 \pm 0,039 \pm 0,459 / {}^{+0,459}_{-0,972}\right) \times 10^{-15}$
7,075–7,1675	$\left(1,238 \pm 0,024 \pm 0,209 / {}^{+0,209}_{-0,397}\right) \times 10^{-15}$
7,1675–7,26	$\left(6,576 \pm 0,152 \pm 1,132 / {}^{+1,145}_{-2,186}\right) \times 10^{-16}$
7,26–7,3525	$\left(3,929 \pm 0,101 \pm 0,684 / {}^{+1,175}_{-1,275}\right) \times 10^{-16}$
7,3525–7,445	$\left(2,296 \pm 0,068 \pm 0,409 / {}^{+0,409}_{-0,913}\right) \times 10^{-16}$
7,445–7,5375	$\left(1,096 \pm 0,041 \pm 0,201 / {}^{+0,230}_{-0,491}\right) \times 10^{-16}$
7,5375–7,63	$\left(4,602 \pm 0,265 \pm 0,978 / {}^{+1,300}_{-0,978}\right) \times 10^{-17}$
7,63–7,7225	$\left(2,818 \pm 0,199 \pm 0,628 / {}^{+1,337}_{-1,016}\right) \times 10^{-17}$
7,7225–7,815	$\left(1,762 \pm 0,129 \pm 0,371 / {}^{+0,371}_{-0,960}\right) \times 10^{-17}$
7,815–7,9075	$\left(9,231 \pm 0,891 \pm 1,991 / {}^{+1,991}_{-2,557}\right) \times 10^{-18}$
7,9075–8,0	$\left(4,788 \pm 0,475 \pm 1,001 / {}^{+1,001}_{-3,306}\right) \times 10^{-18}$

Таблица 5 — Полученные в данном исследовании значения дифференциального потока протонов (dJ_p/dE , $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$dJ_p/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$
6,15–6,2425	$\left(3,105 \pm 0,010 \pm 0,528 / {}^{+1,474}_{-1,147}\right) \times 10^{-13}$
6,2425–6,335	$\left(2,152 \pm 0,007 \pm 0,348 / {}^{+0,682}_{-0,759}\right) \times 10^{-13}$
6,335–6,4275	$\left(1,336 \pm 0,005 \pm 0,214 / {}^{+0,522}_{-0,418}\right) \times 10^{-13}$
6,4275–6,52	$\left(7,307 \pm 0,034 \pm 1,164 / {}^{+2,905}_{-2,340}\right) \times 10^{-14}$
6,52–6,6125	$\left(3,610 \pm 0,019 \pm 0,565 / {}^{+1,654}_{-1,055}\right) \times 10^{-14}$
6,6125–6,705	$\left(2,082 \pm 0,013 \pm 0,329 / {}^{+0,990}_{-0,625}\right) \times 10^{-14}$
6,705–6,7975	$\left(1,093 \pm 0,009 \pm 0,174 / {}^{+0,527}_{-0,337}\right) \times 10^{-14}$
6,7975–6,89	$\left(5,204 \pm 0,053 \pm 0,878 / {}^{+2,701}_{-1,646}\right) \times 10^{-15}$
6,89–6,9825	$\left(2,558 \pm 0,034 \pm 0,484 / {}^{+1,418}_{-0,845}\right) \times 10^{-15}$
6,9825–7,075	$\left(1,207 \pm 0,020 \pm 0,256 / {}^{+0,686}_{-0,418}\right) \times 10^{-15}$
7,075–7,1675	$\left(3,916 \pm 0,080 \pm 0,920 / {}^{+3,448}_{-0,920}\right) \times 10^{-16}$
7,1675–7,26	$\left(1,837 \pm 0,050 \pm 0,454 / {}^{+2,063}_{-0,454}\right) \times 10^{-16}$
7,26–7,3525	$\left(9,385 \pm 0,302 \pm 2,289 / {}^{+9,115}_{-2,289}\right) \times 10^{-17}$
7,3525–7,445	$\left(5,316 \pm 0,257 \pm 1,430 / {}^{+6,727}_{-1,430}\right) \times 10^{-17}$
7,445–7,5375	$\left(2,181 \pm 0,118 \pm 0,607 / {}^{+1,657}_{-0,607}\right) \times 10^{-17}$
7,5375–7,63	$\left(7,364 \pm 0,546 \pm 2,333 / {}^{+9,173}_{-2,333}\right) \times 10^{-18}$
7,63–7,7225	$\left(4,285 \pm 0,458 \pm 1,510 / {}^{+5,566}_{-1,514}\right) \times 10^{-18}$
7,7225–7,815	$\left(3,000 \pm 0,356 \pm 0,925 / {}^{+2,947}_{-0,925}\right) \times 10^{-18}$
7,815–7,9075	$\left(1,312 \pm 0,178 \pm 0,434 / {}^{+0,837}_{-0,434}\right) \times 10^{-18}$
7,9075–8,0	$\left(6,995 \pm 1,096 \pm 2,247 / {}^{+7,120}_{-2,247}\right) \times 10^{-19}$

Таблица 6 — Полученные в данном исследовании значения дифференциального потока гелия (dJ_{He}/dE , $[m^{-2}c^{-1}sr^{-1}GeV^{-1}]$) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/GeV)$	$dJ_{He}/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ $[m^{-2}c^{-1}sr^{-1}GeV^{-1}]$
6,15–6,2425	$(1,037 \pm 0,012 \pm 0,253 / {}^{+0,456}_{-0,883}) \times 10^{-13}$
6,2425–6,335	$(3,724 \pm 0,099 \pm 2,056 / {}^{+5,199}_{-2,056}) \times 10^{-14}$
6,335–6,4275	$(2,898 \pm 0,060 \pm 0,742 / {}^{+1,980}_{-1,978}) \times 10^{-14}$
6,4275–6,52	$(1,819 \pm 0,038 \pm 0,405 / {}^{+1,143}_{-1,212}) \times 10^{-14}$
6,52–6,6125	$(8,963 \pm 0,211 \pm 1,981 / {}^{+5,800}_{-7,387}) \times 10^{-15}$
6,6125–6,705	$(5,234 \pm 0,138 \pm 1,079 / {}^{+3,545}_{-5,831}) \times 10^{-15}$
6,705–6,7975	$(3,064 \pm 0,092 \pm 0,591 / {}^{+1,918}_{-3,088}) \times 10^{-15}$
6,7975–6,89	$(2,017 \pm 0,058 \pm 0,378 / {}^{+0,876}_{-1,559}) \times 10^{-15}$
6,89–6,9825	$(1,264 \pm 0,038 \pm 0,250 / {}^{+0,417}_{-0,780}) \times 10^{-15}$
6,9825–7,075	$(7,069 \pm 0,221 \pm 1,438 / {}^{+1,905}_{-3,173}) \times 10^{-16}$
7,075–7,1675	$(3,308 \pm 0,077 \pm 0,693 / {}^{+0,693}_{-1,653}) \times 10^{-16}$
7,1675–7,26	$(1,633 \pm 0,048 \pm 0,324 / {}^{+0,324}_{-0,549}) \times 10^{-16}$
7,26–7,3525	$(9,485 \pm 0,281 \pm 1,789 / {}^{+1,975}_{-2,752}) \times 10^{-17}$
7,3525–7,445	$(4,490 \pm 0,276 \pm 1,085 / {}^{+1,085}_{-2,570}) \times 10^{-17}$
7,445–7,5375	$(2,152 \pm 0,112 \pm 0,478 / {}^{+1,042}_{-0,661}) \times 10^{-17}$
7,5375–7,63	$(8,341 \pm 0,546 \pm 2,179 / {}^{+4,299}_{-2,840}) \times 10^{-18}$
7,63–7,7225	$(3,475 \pm 0,445 \pm 1,281 / {}^{+3,919}_{-1,281}) \times 10^{-18}$
7,7225–7,815	$(2,112 \pm 0,317 \pm 0,644 / {}^{+1,978}_{-0,804}) \times 10^{-18}$
7,815–7,9075	$(1,364 \pm 0,161 \pm 0,410 / {}^{+1,230}_{-0,477}) \times 10^{-18}$
7,9075–8,0	$(6,873 \pm 0,926 \pm 1,927 / {}^{+3,856}_{-2,679}) \times 10^{-19}$

Таблица 7 — Полученные в данном исследовании значения дифференциального потока углерода (dJ_C/dE , $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$dJ_C/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$
6,15–6,2425	$(6,513 \pm 0,064 \pm 1,544 / {}^{+1,894}_{-4,507}) \times 10^{-14}$
6,2425–6,335	$(6,701 \pm 0,128 \pm 2,424 / {}^{+2,424}_{-4,709}) \times 10^{-14}$
6,335–6,4275	$(2,691 \pm 0,050 \pm 0,607 / {}^{+0,607}_{-1,865}) \times 10^{-14}$
6,4275–6,52	$(1,293 \pm 0,026 \pm 0,270 / {}^{+0,349}_{-0,904}) \times 10^{-14}$
6,52–6,6125	$(7,357 \pm 0,172 \pm 1,493 / {}^{+1,493}_{-4,939}) \times 10^{-15}$
6,6125–6,705	$(3,814 \pm 0,103 \pm 0,735 / {}^{+0,735}_{-2,248}) \times 10^{-15}$
6,705–6,7975	$(1,930 \pm 0,063 \pm 0,375 / {}^{+0,375}_{-1,042}) \times 10^{-15}$
6,7975–6,89	$(1,125 \pm 0,036 \pm 0,218 / {}^{+0,218}_{-0,548}) \times 10^{-15}$
6,89–6,9825	$(6,811 \pm 0,224 \pm 1,399 / {}^{+1,399}_{-3,189}) \times 10^{-16}$
6,9825–7,075	$(4,218 \pm 0,135 \pm 0,869 / {}^{+0,869}_{-1,899}) \times 10^{-16}$
7,075–7,1675	$(2,580 \pm 0,066 \pm 0,568 / {}^{+0,568}_{-1,101}) \times 10^{-16}$
7,1675–7,26	$(1,437 \pm 0,045 \pm 0,319 / {}^{+0,319}_{-0,998}) \times 10^{-16}$
7,26–7,3525	$(8,669 \pm 0,288 \pm 1,881 / {}^{+1,881}_{-3,775}) \times 10^{-17}$
7,3525–7,445	$(5,732 \pm 0,287 \pm 1,386 / {}^{+1,386}_{-2,797}) \times 10^{-17}$
7,445–7,5375	$(2,570 \pm 0,130 \pm 0,583 / {}^{+0,583}_{-1,214}) \times 10^{-17}$
7,5375–7,63	$(1,138 \pm 0,069 \pm 0,281 / {}^{+0,281}_{-0,562}) \times 10^{-17}$
7,63–7,7225	$(7,481 \pm 0,650 \pm 2,013 / {}^{+2,013}_{-4,076}) \times 10^{-18}$
7,7225–7,815	$(3,607 \pm 0,420 \pm 0,939 / {}^{+2,174}_{-1,937}) \times 10^{-18}$
7,815–7,9075	$(2,271 \pm 0,237 \pm 0,585 / {}^{+1,168}_{-1,186}) \times 10^{-18}$
7,9075–8,0	$(9,971 \pm 1,282 \pm 2,534 / {}^{+7,396}_{-5,288}) \times 10^{-19}$

Таблица 8 — Полученные в данном исследовании значения дифференциального потока кремния (dJ_{Si}/dE , $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$dJ_{Si}/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ $[\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}]$
6,15–6,2425	$(6,735 \pm 0,089 \pm 1,654 / {}^{+3,818}_{-2,247}) \times 10^{-14}$
6,2425–6,335	$(2,497 \pm 0,078 \pm 1,592 / {}^{+3,170}_{-1,592}) \times 10^{-14}$
6,335–6,4275	$(2,276 \pm 0,047 \pm 0,481 / {}^{+1,195}_{-0,481}) \times 10^{-14}$
6,4275–6,52	$(1,091 \pm 0,027 \pm 0,221 / {}^{+0,554}_{-0,223}) \times 10^{-14}$
6,52–6,6125	$(6,738 \pm 0,157 \pm 1,234 / {}^{+2,757}_{-2,938}) \times 10^{-15}$
6,6125–6,705	$(3,893 \pm 0,104 \pm 0,696 / {}^{+0,719}_{-1,654}) \times 10^{-15}$
6,705–6,7975	$(2,244 \pm 0,068 \pm 0,409 / {}^{+0,409}_{-0,993}) \times 10^{-15}$
6,7975–6,89	$(1,028 \pm 0,039 \pm 0,207 / {}^{+0,207}_{-0,405}) \times 10^{-15}$
6,89–6,9825	$(5,513 \pm 0,241 \pm 1,232 / {}^{+1,232}_{-2,451}) \times 10^{-16}$
6,9825–7,075	$(3,097 \pm 0,142 \pm 0,695 / {}^{+0,695}_{-1,350}) \times 10^{-16}$
7,075–7,1675	$(1,560 \pm 0,053 \pm 0,346 / {}^{+0,498}_{-0,475}) \times 10^{-16}$
7,1675–7,26	$(1,037 \pm 0,040 \pm 0,240 / {}^{+0,240}_{-0,421}) \times 10^{-16}$
7,26–7,3525	$(6,769 \pm 0,252 \pm 1,533 / {}^{+1,533}_{-3,241}) \times 10^{-17}$
7,3525–7,445	$(4,381 \pm 0,283 \pm 1,124 / {}^{+2,379}_{-1,573}) \times 10^{-17}$
7,445–7,5375	$(2,277 \pm 0,126 \pm 0,549 / {}^{+0,549}_{-0,931}) \times 10^{-17}$
7,5375–7,63	$(1,077 \pm 0,069 \pm 0,291 / {}^{+0,291}_{-0,471}) \times 10^{-17}$
7,63–7,7225	$(7,968 \pm 0,754 \pm 2,301 / {}^{+2,301}_{-4,024}) \times 10^{-18}$
7,7225–7,815	$(5,482 \pm 0,570 \pm 1,464 / {}^{+1,464}_{-2,665}) \times 10^{-18}$
7,815–7,9075	$(2,426 \pm 0,249 \pm 0,633 / {}^{+0,633}_{-1,091}) \times 10^{-18}$
7,9075–8,0	$(1,300 \pm 0,149 \pm 0,322 / {}^{+0,322}_{-0,569}) \times 10^{-18}$

Таблица 9 — Полученные в данном исследовании значения дифференциального потока кремния (dJ_{Fe}/dE , [$\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}$]) в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$dJ_{Fe}/dE \pm \text{Стат.} \pm \text{Баз.} / \pm \text{Теор.},$ [$\text{м}^{-2}\text{с}^{-1}\text{ср}^{-1}\text{ГэВ}^{-1}$]
6,15–6,2425	$(2,711 \pm 0.059 \pm 1.007 / {}^{+6,223}_{-2,029}) \times 10^{-14}$
6,2425–6,335	$(1,503 \pm 0.043 \pm 0.712 / {}^{+3,356}_{-1,223}) \times 10^{-14}$
6,335–6,4275	$(6,606 \pm 0.217 \pm 2.079 / {}^{+19,007}_{-5,117}) \times 10^{-15}$
6,4275–6,52	$(3,711 \pm 0.128 \pm 1.089 / {}^{+9,614}_{-2,575}) \times 10^{-15}$
6,52–6,6125	$(1,573 \pm 0.073 \pm 0.408 / {}^{+6,052}_{-1,589}) \times 10^{-15}$
6,6125–6,705	$(8,312 \pm 0.468 \pm 2.201 / {}^{+34,238}_{-3,911}) \times 10^{-16}$
6,705–6,7975	$(3,797 \pm 0.287 \pm 1.212 / {}^{+19,159}_{-2,139}) \times 10^{-16}$
6,7975–6,89	$(3,045 \pm 0.192 \pm 0.869 / {}^{+9,794}_{-1,064}) \times 10^{-16}$
6,89–6,9825	$(2,124 \pm 0.131 \pm 0.641 / {}^{+5,707}_{-0,641}) \times 10^{-16}$
6,9825–7,075	$(1,418 \pm 0.086 \pm 0.421 / {}^{+3,379}_{-0,421}) \times 10^{-16}$
7,075–7,1675	$(1,015 \pm 0.042 \pm 0.285 / {}^{+1,455}_{-0,792}) \times 10^{-16}$
7,1675–7,26	$(6,321 \pm 0.299 \pm 1.791 / {}^{+9,239}_{-3,868}) \times 10^{-17}$
7,26–7,3525	$(4,984 \pm 0.225 \pm 1.350 / {}^{+6,421}_{-2,428}) \times 10^{-17}$
7,3525–7,445	$(3,040 \pm 0.196 \pm 0.897 / {}^{+4,124}_{-2,829}) \times 10^{-17}$
7,445–7,5375	$(1,777 \pm 0.109 \pm 0.514 / {}^{+2,189}_{-0,881}) \times 10^{-17}$
7,5375–7,63	$(8,163 \pm 0.585 \pm 2.629 / {}^{+10,512}_{-5,203}) \times 10^{-18}$
7,63–7,7225	$(4,974 \pm 0.514 \pm 1.675 / {}^{+7,294}_{-2,984}) \times 10^{-18}$
7,7225–7,815	$(3,423 \pm 0.398 \pm 0.986 / {}^{+4,802}_{-3,178}) \times 10^{-18}$
7,815–7,9075	$(1,858 \pm 0.219 \pm 0.540 / {}^{+2,196}_{-1,146}) \times 10^{-18}$
7,9075–8,0	$(1,104 \pm 0.143 \pm 0.300 / {}^{+1,198}_{-0,819}) \times 10^{-18}$

Таблица 10 — Полученные в данном исследовании значения $\langle \ln A \rangle$ в зависимости от энергии первичной частицы

Энерг. интервал, $\log_{10}(E/\text{ГэВ})$	$\langle \ln A \rangle$	\pm Стат. неопр.	\pm «Базовая» неопр.	– Теор. неопр. +
6,15–6,2425	1,114	0,008	0,117	0,385, 0,526
6,2425–6,335	1,007	0,013	0,138	0,336, 0,538
6,335–6,4275	0,957	0,011	0,094	0,317, 0,448
6,4275–6,52	0,914	0,011	0,086	0,311, 0,443
6,52–6,6125	0,980	0,013	0,088	0,374, 0,384
6,6125–6,705	0,955	0,015	0,087	0,357, 0,388
6,705–6,7975	0,973	0,017	0,092	0,360, 0,408
6,7975–6,89	1,058	0,020	0,120	0,378, 0,450
6,89–6,9825	1,165	0,023	0,164	0,409, 0,507
6,9825–7,075	1,303	0,027	0,196	0,425, 0,547
7,075–7,1675	1,638	0,025	0,249	0,587, 0,360
7,1675–7,26	1,799	0,034	0,274	0,718, 0,416
7,26–7,3525	1,967	0,038	0,285	0,656, 0,448
7,3525–7,445	2,060	0,064	0,326	0,757, 0,517
7,445–7,5375	2,200	0,064	0,338	0,606, 0,547
7,5375–7,63	2,359	0,082	0,393	0,760, 0,639
7,63–7,7225	2,482	0,131	0,435	0,828, 0,718
7,7225–7,815	2,492	0,155	0,374	0,787, 0,637
7,815–7,9075	2,502	0,148	0,366	0,639, 0,619
7,9075–8,0	2,549	0,174	0,360	0,741, 0,613