

Курс «Нейронные сети и их применение в научных исследованиях»

Метод разделения широких атмосферных ливней по первичной массе с помощью машинного обучения для черенковского телескопа типа СФЕРА

В.С. Латыпова

Московский государственный университет имени М.В. Ломоносова, физический факультет, кафедра физики космоса. Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2.

Аннотация

Важной задачей физики высоких энергий является нахождение массового состава космических лучей в диапазоне энергий 1–100 ПэВ. Целью эксперимента СФЕРА-2 является определение массового состава первичного космического излучения. Чем лучше происходит разделение событий, вызванных черенковским светом широких атмосферных ливней, тем точнее происходит оценка средней массы по выборке.

В настоящей работе разработан эффективный способ разделения первичных ядер, образовавших широкие атмосферные ливни, на основе смоделированных событий аппарата СФЕРА с помощью методов машинного обучения. Поскольку для моделирования искусственных событий используется модель ядро-ядерного взаимодействия, результат разделения событий, а, следовательно, и оценка средней массы, может меняться. В эксперименте СФЕРА-2 эта проблема решена. Во-первых, благодаря использованию данных о черенковском свете, которые мало зависят от модели адронного взаимодействия. Во-вторых, обучение нейронной сети произведено одновременно на двух моделях ядерного взаимодействия, которые сильно отличаются между собой в диапазоне энергий 1–100 ПэВ. Это модели *QGSJET-01* и *QGSJETII-04*. Таким образом, обеспечена независимость обработки экспериментальных данных от выбора модели ядерного взаимодействия.

Методами машинного обучения решена задача регрессии. Разделение событий на три группы ядер – протонов (p), азота (N), железа (Fe) с помощью нейронной сети происходит более точно, чем при использовании традиционных методов.

§1. Введение

Исследование состава космического излучения сверхвысоких энергий ($> 10^{15}$ эВ) является одной из актуальных задач физики космических лучей. В связи с малой интенсивностью первичных частиц с энергией выше 1 ПэВ (~ 1 частица на м^2 на ср в год), их изучение происходит косвенными методами.

Наиболее распространенным методом нахождения параметров первичных частиц (таких как энергетический спектр, массовый состав) является измерение различных компонент широких атмосферных ливней (ШАЛ). Например, электронно-фотонной, мюонной, адронной компоненты ШАЛ, флуоресцентного излучения, радиоизлучения, акустических эффектов, излучения Вавилова-Черенкова.

Чаще всего используются наземные установки, например, HiRes [1], TAIGA [2], Telescope Array (TA) [3], Якутская установка [4, 5] и другие. Однако проводился и воздушный эксперимент СФЕРА-2 [6], который использовал для регистрации ШАЛ оптическое излучение Вавилова-Черенкова, также называемое черенковским светом, отраженное от заснеженной поверхности льда озера Байкал.

Более 50 лет назад было обнаружено изменение наклона энергетического спектра первичных космических лучей (ПКЛ) при энергиях около 3 ПэВ. Причину возникновения изменения крутизны спектра связывают с изменением состава ПКЛ, который, в свою очередь, способен дать важную информацию о механизмах ускорения и устройстве источников космических лучей. Вопрос о составе первичных космических лучей сверхвысоких энергий в настоящее время далек от окончательного решения [7, 8, 9, 10]. Экспериментальной группе KASCADE-Grande удалось выделить спектры двух групп ядер [11]. Более детальной классификации мешает неопределенность модели сильного взаимодействия. Результаты различных экспериментов по среднему логарифму массового числа в области энергий 10 ПэВ отличаются в несколько раз [12].

Метод проекта СФЕРА основан на идее А.Е. Чудакова регистрировать черенковский свет ШАЛ, отраженный от поверхности земли, фотоприемниками, установленными на летательных аппаратах [13]. Детектор СФЕРА-2 регистрирует время и место взаимодействия отраженных фотонов с фотоэлектронными умножителями.

Восстановление свойств космических лучей по атмосферным ливням включает направление прихода, энергию и массовый состав. В эксперименте СФЕРА-2 угол падения оси ливня извлекается по временным задержкам, энергия восстанавливается по суммарным количествам фотонов в одном зарегистрированном ливне. Восстановление средней массы традиционными методами происходит с помощью критериев, основанных на функциях пространственного распределения черенковского света атмосферных ливней [14]. Восстановление массы тем точнее, чем лучше критерий разделяет события по первичным массам. В качестве альтернативы критерию в настоящей работе представлено решение задачи регрессии методом глубокого обучения для эффективного разделения событий атмосферных ливней по первичным массам из смоделированных данных телескопа СФЕРА-2.

§2. Моделирование сигнала ШАЛ

Для ливней с небольшими зенитными углами ($0 - 20^\circ$) было произведено полное моделирование сигнала ШАЛ в телескопе СФЕРА-2.

С помощью программы CORSIKA7.56 [15] проведено моделирование пространственно-временного распределения черенковского света на уровне снега. Для экономии машинного времени в код CORSIKA закладываются характеристики ФЭУ и прослеживаются только те черенковские фотоны ШАЛ, которые с вероятностью единица выбьют фотоэлектроны из фотокатода. Результатом работы программы является количество «фотоэлектронов» в ячейках $2.5 \times 2.5 \text{ м}^2$ от данной первичной частицы с заданными первичными параметрами. Эти ячейки образуют пятно черенковского света площадью $3200 \times 3200 \text{ м}^2$ с центром на оси ливня. На снегу диаметр пятна черенковского света, который содержит полезную информацию, составляет от 1000 до 2000 м.

Из распределения на снегу генерируются «фотоэлектроны» на входном окне телескопа. Поскольку свет ШАЛ диффузно отражается от снега, то в детектор попадает лишь малая его часть. На этом этапе остаются лишь те «фотоэлектроны», которые имеют шанс попасть в пиксели мозаики.

С помощью программы GEANT4 происходит прослеживание «фотоэлектронов» от входного окна телескопа до пикселей мозаики с отражением от зеркала и поглощением на бленде (рис. 1). Радиус мозаики ФЭУ составляет 220 мм.

На выходе моделирование дает ту же информацию, что и воздушный детектор. Аэростатная установка СФЕРА-2 (рис. 1) состоит из сферического зеркала, в фокусе которого установлена мозаика из 109 фотоэлектронных умножителей (ФЭУ). Каждый из них обозревает некоторую площадь на заснеженной поверхности льда озера Байкал.

Разделение событий традиционным методом и методом глубокого обучения используют один и тот же набор моделирования космических явлений по Монте-Карло. Было выбрано две модели ядро-ядерного взаимодействия, QGSJET01 [16], QGSJETII-04 [17], которые сильно различаются меж-

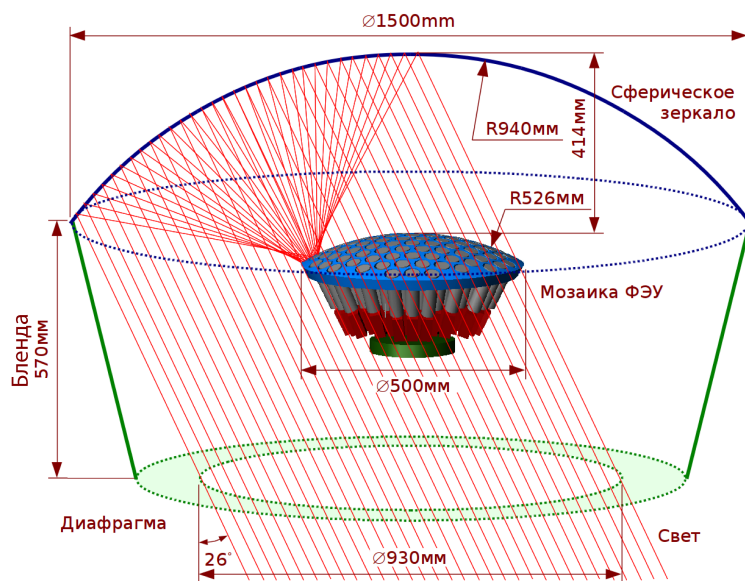


Рисунок 1. Схема детектора СФЕРА-2

ду собой в диапазоне энергий 1–100 ПэВ [18]. Выбрана атмосфера №1 в CORSIKA7 — аппроксимация Линсли стандартной американской атмосферы. Энергия первичной частицы составила 10 ПэВ, угол наклона оси ливня $10 - 20^\circ$, высота расположения детектора над уровнем снежной поверхности 900 метров. Тогда как заснеженная поверхность озера Байкал находится на 455 м над уровнем моря.

Чтобы более эффективно использовать доступные ливни CORSIKA, каждый ливень копируется 100 раз и выбрасывается в случайные места в пределах круга радиусом R с центром на оси детектора. Таким образом, для каждого из 5 ядер (p, He, N, S, Fe) смоделировано по 6000 событий атмосферных ливней.

§3. Предобработка данных

Входные данные нейронной сети состоят из видеоряда и двух двумерных карт нормализованных времен прихода ливня и суммарных сигналов. Данные приведены к размерности (1, 50, 11, 11), где размер 11×11 это сетка ФЭУ, в ячейках которой записано количество фотонов, попавших в определенный ФЭУ за 15 нс. Размерность 50 отвечает за время регистрации событий. Разбивая каждое событие на 15 нс можно наблюдать от 60 до 90 кадров. Но только в первых 50 кадрах срабатывают больше, чем 5 ФЭУ. Чтобы избавиться от большого числа нулей было оставлено только 50 кадров с момента попадания первого фотона.

На вход сети в дополнение к видеоряду подается две двумерные карты. На рис. 2а) представлен пример карты нормализованных времен прихода ливня (рассчитывается среднее время прихода фотонов в определенный ФЭУ) для одного события. Белые точки определяют центры расположения ФЭУ на мозике телескопа СФЕРА-2. На рис. 2б) показана карта суммарных сигналов (рассчитывается число фотонов, попавших в каждый ФЭУ за все время регистрации одного события) атмосферного ливня.

Если в одном событии регистрировалось менее 1200 фотоэлектронов, то такое событие не рассматривалось, поскольку в экспериментальных данных нет событий с меньшим числом фотоэлектронов. После установки данного ограничения из 6000 событий осталось около 3700 ± 500 событий для каждого ядра. Это означает, что существенного дисбаланса в данных нет.

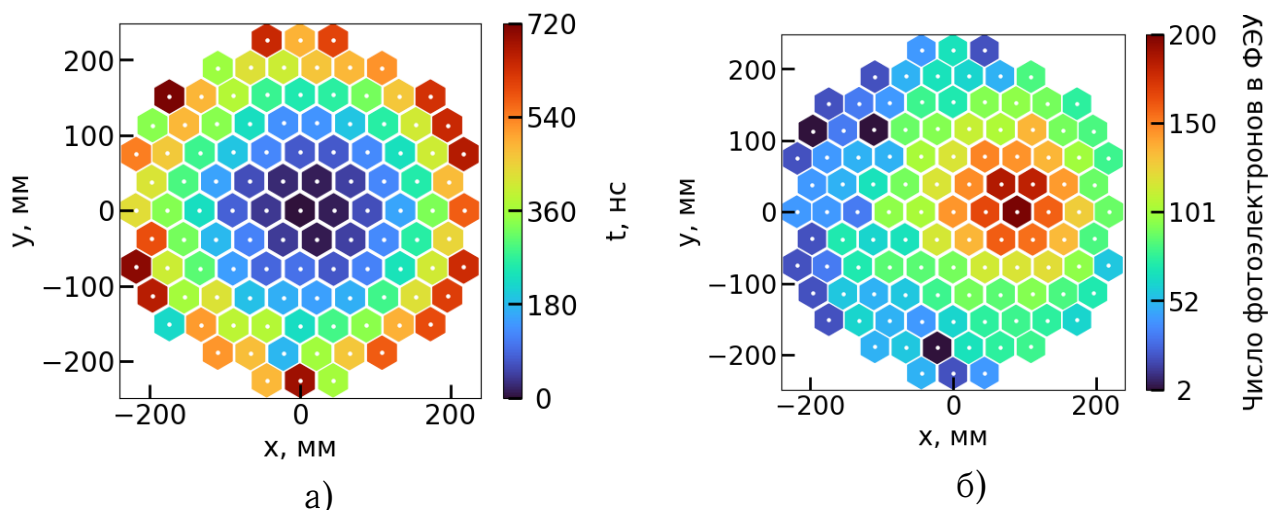


Рисунок 2. а) Карта среднего времени активации ФЭУ; б) Карта суммарного числа фотонов в ФЭУ (форма пространственного распределения)

§4. Поиск архитектуры

Pierre Auger Observatory – одна из крупнейших в мире обсерваторий космических лучей, состоящая из двух независимых типов детекторов: 27 флуоресцентных телескопов и 1660 водно-черенковских детекторов, разнесенных на большие расстояния и занимающие площадь порядка 3000 кв. км.

Одной из задач данного эксперимента является исследование массового состава космических лучей с энергиями выше 0.3×10^{18} эВ. Для его изучения создали нейронную сеть AixNet [19], которую использовали для получения информации о параметре ШАЛ, X_{max} – глубине максимума развития атмосферного ливня. Параметр X_{max} связан с массой первичной частицы. Сеть AixNet предназначена для обработки характеристик широких атмосферных ливней.

Поскольку данные обсерватории Pierre Auger схожи по типу, но не по физическому смыслу, с данными эксперимента СФЕРА, были рассмотрены нейронные сети, созданные разными группами обсерватории Pierre Auger.

В статье [20] применяют глубокие нейронные сети для оценки содержания мюонов в ливнях при измерении водными черенковскими детекторами обсерватории Pierre Auger. Обучают нейронные сети только на данных с поверхностных детекторов смешанной выборкой смоделированных событий, но лишь на одной модели адронного взаимодействия.

§5. Архитектура нейронной сети

В ходе работы было опробовано несколько архитектур нейронных сетей. Поиск лучшей архитектуры проходил по решению задачи классификации. Одной из лучших оказалась сеть AixNet [19]. Поэтому она была модифицирована под эксперимент СФЕРА–2. Архитектура сети, изображенная на рис. 3, состоит из двух функциональных частей. Первая часть характеризует временные треки. Затем происходит объединение карт времени прихода, общего сигнала и извлеченных признаков временного трека. Последний из двух полносвязных слоёв выводит логарифм массы первичной частицы, образовавшей атмосферный ливень, если решается задача регрессии, или выводит тип первичной частицы при задаче классификации. В качестве фреймворка для глубокого обучения использован Pytorch.

Характеристика временного трека. Первая часть сети принимает в качестве входных данных тензор размерности (1, 50, 11, 11). Посколь-

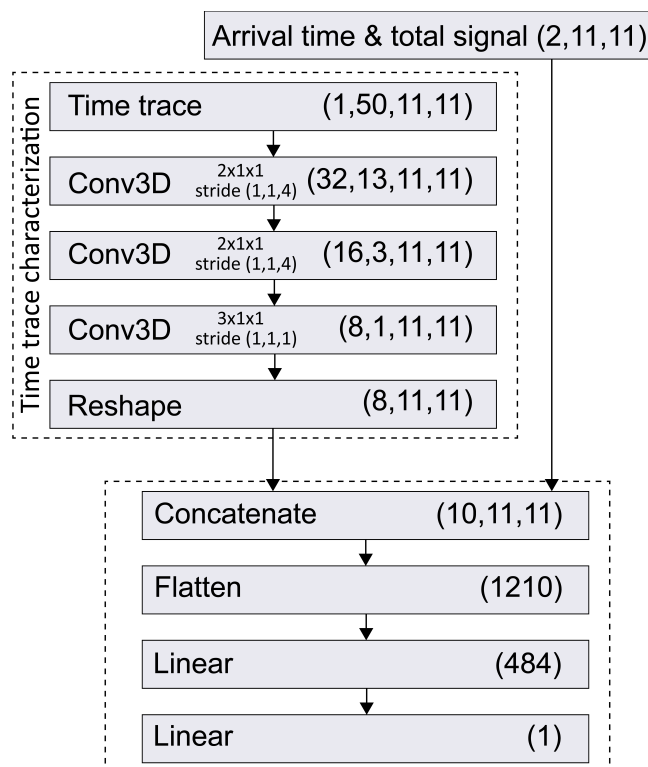


Рисунок 3. Архитектура нейронной сети

ку всего 109 ФЭУ, создана сетка 11×11 , где индекс соответствует номеру ФЭУ, а остальное заполнено нулями. Тензор содержит временные треки по 50 временным интервалам для каждого ФЭУ. Первый размер тензора 1 требуется по техническим причинам. Чтобы охарактеризовать эти временные треки, применяются три последовательных слоя одномерных сверток к каждому ФЭУ отдельно. Технически это реализуется с помощью 3D-сверток с фильтрами размером $\{2 \times 1 \times 1; 2 \times 1 \times 1; 3 \times 1 \times 1\}$. Для трех слоев использовано количество фильтров $\{32, 16, 8\}$, скользящие по временным трекам с шагом $\{4, 4, 1\}$ без заполнения. Шагая, каждая свертка видит временной трек с более низким разрешением и более широким масштабом. После третьей свертки временное измерение пропадает, и выходной тензор имеет размерность $(8, 1, 11, 11)$ с первым измерением, содержащим 8 извлеченных признаков.

Измерение необходимое для времени удаляется путем изменения размерности до $(8, 11, 11)$. Эту часть сети можно интерпретировать как единую функцию, характеризующую временной трек каждого ФЭУ небольшим числом признаков. Так как вся сеть обучается как единое целое, функция научится извлекать те 8 признаков, которые наиболее полезны для следующего этапа сети. Изображения извлеченных признаков временной трассы объединяются со временем прихода и общим сигналом, что дает тензор размерности $(10, 11, 11)$.

Выходной слой. Выходной тензор $(10, 11, 11)$ выравнивается и подается в первый полносвязный слой. Второй полносвязный слой без функции активации выводит логарифм массы первичной частицы, образовавшей атмосферный ливень, если решается задача регрессии. Так, выход нейронной сети дает альтернативный критерий, который способен разделять события по первичным массам.

Дизайн сети. Использована функция активации ReLU в качестве нелинейности после всех взвешенных слоев, кроме полносвязных слоев. Протестированные варианты сети включают в себя простую серию сверток и разделимые свертки (SepConv2D как в статье [21]), которые имели разную глубину и количество параметров. Ни один из этих вариантов не показал улучшенных результатов. Тесты также не показали преимуществ использования нормализации сигналов ни в одной части сети.

Обучение. Задача поиска альтернативного критерия для разделения событий по массам поставлена как проблема регрессии с использованием среднеквадратичной ошибки между прогнозируемым и целевым значением в качестве функции потерь. В качестве оптимизатора использован ADAM [22] со стандартными значениями начальной скорости обучения $\alpha = 10^{-3}$ и скоростей экспоненциального затухания первых двух моментов $\beta_1 = 0.9$, $\beta_2 = 0.999$. Сеть имеет 484 параметра для восстановления логарифма массы. Обучение происходит на $\sim 20\,000$ событий. Время обучения на эпоху составляет ~ 4 с.

§6. Результат классификации

Для отладки работы сети и ее оптимизации была решена задача классификации трех ядер (протонов, ядер азота и ядер железа). На рис. 4 отображена вероятность определения ядер. Цветом обозначено истинное ядро, а по оси абсцисс предсказываемое. Левее от подписей оси ОХ показаны результаты традиционного метода, правее результаты классификации нейронной сети. На 20% лучше стали определяться протоны и ядра азота.

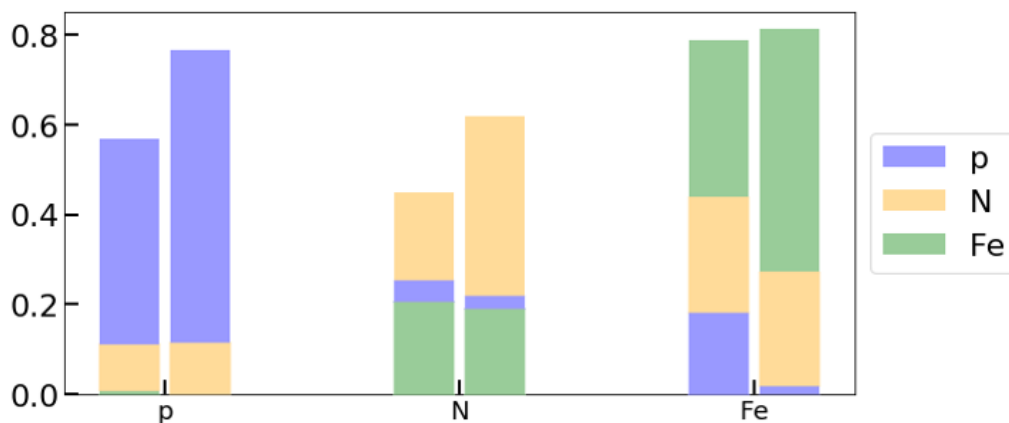


Рисунок 4. Вероятность определения ядер. Цветом обозначено истинное ядро, а по оси абсцисс предсказываемое. Левее от подписей оси ОХ показаны результаты традиционного метода, правее результаты классификации нейронной сети

Результат показан на тестовой выборке. Для каждого ядра в тестовой выборке порядка 1 500 событий. На вход модели подавались выборки одновременно по двум моделям адронного взаимодействия.

Множество научных групп решают задачу разделения двух ядер, протонов и ядер железа. Команда обсерватории Pierre Auger добилась результатов разделения протонов и ядер железа с ошибкой порядка 10% [23] для одной модели QGSJET. Применив описанную выше нейронную сеть на смоделированных данных СФЕРА-2, ошибки разделения на два ядра составили порядка 4% (рис. 5), при том что модель обучалась одновременно на двух моделях ядерного взаимодействия. Из этого следует, что у метода есть возможность оценивать массы с более высокой точностью.

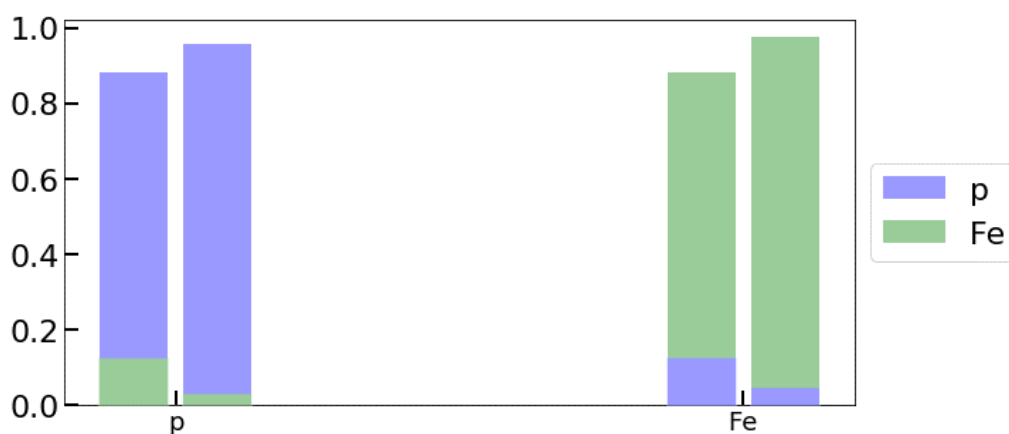


Рисунок 5. Вероятность определения ядер. Цветом обозначено истинное ядро, а по оси абсцисс предсказываемое. Левее от подписей оси ОХ показаны результаты команды Pierre Auger [23], правее результаты классификации нейронной сети

§7. Результат регрессии

7.1. Метрика оценки ошибок разделения событий

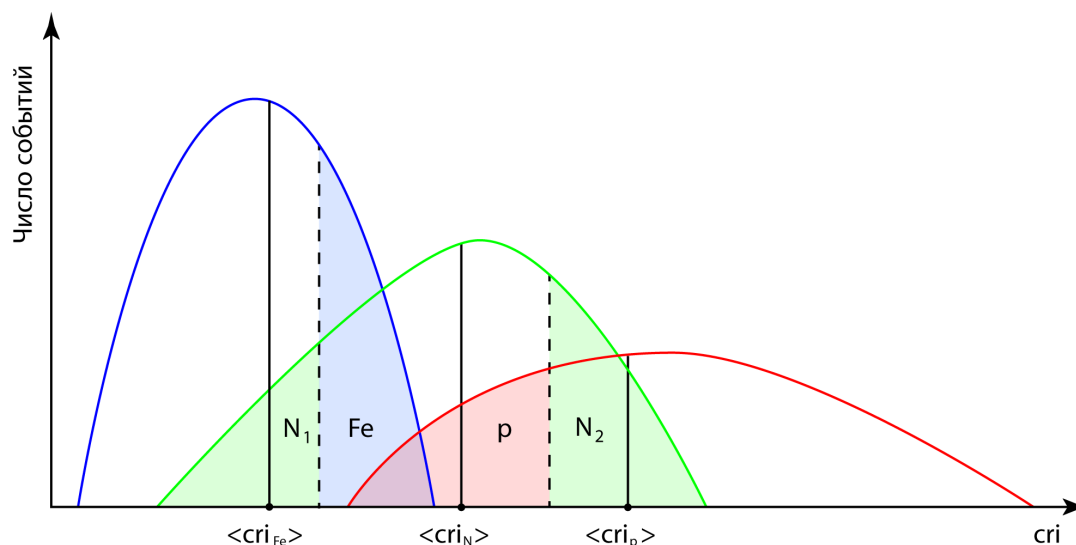


Рисунок 6. Схема разделения ядер. $\langle cri_p \rangle$, $\langle cri_N \rangle$, $\langle cri_{Fe} \rangle$ - средние значения критерия для ядер; N_2 — количество событий от первичных ядер азота, которые лежат правее границы классов p - N ; p — количество событий от первичных протонов, которые лежат левее границы классов p - N . Граница варьируется в пределах области $[\langle cri_N \rangle; \langle cri_p \rangle]$ до получения одинаковых ошибок классификации для классов пары, то есть одинаковых вероятностей $P\{p \rightarrow N\}$ и $P\{N \rightarrow p\}$. Именно такие ошибки классификации приведены в таблице 1. Аналогично для пары N - Fe .

Восстановление средней массы традиционными методами происходит с помощью критериев. Для событий с углом наклона оси ливня меньше 20° построен одномерный критерий, который определяется как отношение интеграла функции пространственного распределения (ФПР) черенковского света атмосферного ливня по области центрального круга радиуса r_1 к интегралу по прилежающему кольцу с внутренним радиусом r_1 и внешним радиусом r_2 (рис. 26) с центром на оси ливня.

После подсчета критериального параметра таким традиционным методом, получается распределение, подобное графику на рис. 6. Рассчитываются математические ожидания для каждого распределения критерия по трем известным массам. Между ними происходит поиск границы (штриховая линия), которая соответствует получению одинаковых ошибок разделения классов пары, то есть одинаковых вероятностей $P\{p \rightarrow N\}$ и $P\{N \rightarrow p\}$. На практике это означает достижение равенства между областями N_2 и p на графике. Где N_2 — это количество событий от первичных ядер азота, которые лежат правее границы классов p — N , область p — это количество событий от первичных протонов, которые лежат левее границы классов p — N . Таким образом, показателем качества разделения по массам является отношение области p к сумме всех событий протонов. Аналогичные ошибки классификации рассчитываются для пары N — Fe .

7.2. Результат разделения с помощью традиционного критерия

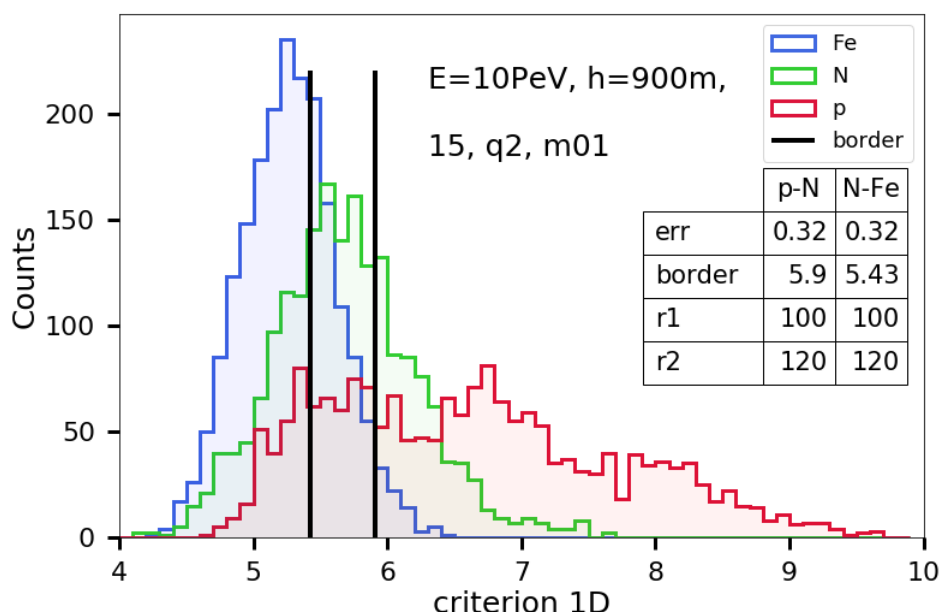


Рисунок 7. Разделение событий атмосферных ливней по трем ядрам традиционным методом с помощью критериев, рассчитанных по ФПР черенковского света ШАЛ; вертикальными линиями обозначены границы разделения

На рисунке 7 приведен пример распределения критериального параметра трех известных ядер для событий с энергией $E = 10$ ПэВ, под углом наклона оси ливня $10 - 20^\circ$, высотой расположения детектора 900 м над снежной поверхностью, модели атмосферы №1 в CORSIKA, модели адронного взаимодействия QGSJETII-04. На графике сплошными вертикальными линиями отмечены границы разделения пар классов. Им соответствуют численные значения критериального параметра, указанные в таблице в строке border. Присущие данным границам ошибки классификации указаны в графе err, и для данного набора первичных параметров ошибка определения протона как азота равна 32%, ошибка определения железа как азота равна 32%.

В таблице 1 приведены результаты разделения классов для разных первичных параметров, высоты расположения детектора и разных моделей адронного взаимодействия на основе расчетов критерия, традиционными методами.

Таблица 1. Лучшие ошибки классификации для энергии первичной частицы $E=10$ ПэВ на основе расчетов критерия, традиционными методами

	$E = 10$ ПэВ, $\theta = 10 - 20^\circ$			
	$h = 500$ м		$h = 900$ м	
	p-N	N-Fe	p-N	N-Fe
QGSJET01	0.36	0.30	0.35	0.26
QGSJETII-04	0.33	0.32	0.32	0.32

7.3. Результат разделения с помощью нейронной сети

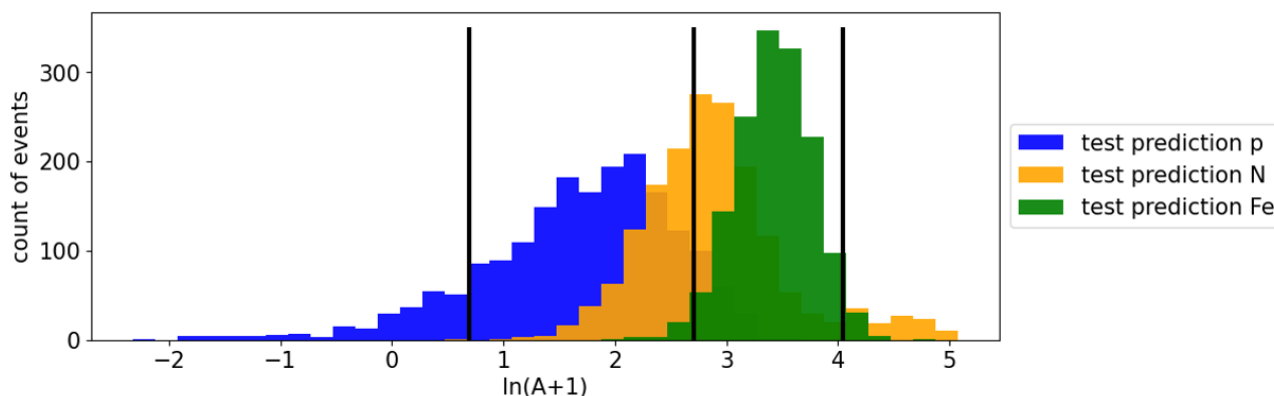


Рисунок 8. Разделение событий атмосферных ливней по трем ядрам с помощью модифицированной нейронной сети AixNet (Sphere AixNet); вертикальными линиями обозначены истинные логарифмы масс

На рис. 8 показаны гистограммы для каждого смоделированного элемента (протон, азот и железо) в натуральном логарифме выхода массы нейронной сети. Ошибки классификации для выхода массы нейронной сети меньше 20%. При этом обучение нейронной сети произведено одновременно на двух моделях ядерного взаимодействия. Разделение событий традиционным методом и методом глубокого обучения используют один и тот же набор моделирования событий. Разделение событий по ядрам с помощью традиционного критерия для каждой модели больше 30%, при том что критерий еще немного меняется в зависимости от модели адронного взаимодействия.

§8. Восстановление массы первичной частицы

На данный момент ЛНС может разгонять частицы до энергий $3.3 \cdot 10^{14}$ эВ [24]. Эти данные способствуют улучшению ядерных моделей, однако при энергиях выше 10^{15} эВ нет возможности сверять теоретические модели с экспериментальными данными. Из-за этого любой подход к определению массы первичной частицы становится зависимым от выбора адронной модели. Целью данной работы является избавление от этой зависимости.

Важной задачей физики космических лучей является нахождение массового состава первичных ядер в диапазоне энергий 1 – 100 ПэВ ($10^{15} - 10^{17}$ эВ). Во многих экспериментах результат оценивания усредненной по событиям массы меняется в зависимости выбора модели ядерного взаимодействия. В эксперименте СФЕРА-2 эта проблема решена, во-первых, благодаря использованию данных о черенковском свете, которые мало зависят от модели адронного взаимодействия. Черенковский свет инициируется ультрарелятивистскими электронами на всем пути прохождения атмосферного ливня, в результате на поверхности видна интегральная картина прохождения ливня в атмосфере. Во-вторых, обучение нейронной сети произведено одновременно на двух моделях ядерного взаимодействия, которые сильно отличаются между собой в диапазоне энергий 1–100 ПэВ [18]. Это модели *QGSJET*–01 и *QGSJETII*–04. Таким образом, обеспечена независимость обработки экспериментальных данных от выбора модели ядерного взаимодействия.

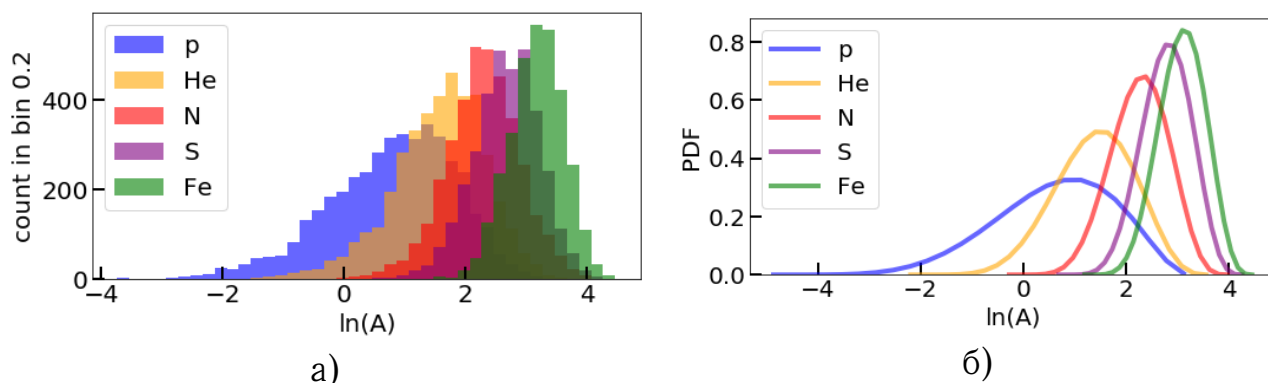


Рисунок 9. а) гистограмма выхода массы нейронной сети по известным ядрам; б) PDF (Probability Density Function) – плотность распределения вероятности выхода массы

Измерение массового состава космических лучей в диапазоне энергий от 1 до 100 ПэВ позволит получить информацию о возникновении «колена» в энергетическом спектре. Все измерения, произведенные на данный момент, имеют большие методические и систематические ошибки. Целью эксперимента СФЕРА-2 является уменьшение всех возможных ошибок процедуры определения массы первичной частицы.

Существует множество способов восстановления массового состава. Один из них разработан для экспериментов СФЕРА для восстановления среднего массового числа [14]. Основные этапы нахождения средней массы:

Формирование смешанной выборки. В начале формируется смешанная выборка, которая состоит из смоделированных событий по нескольким типам первичных частиц. Для каждой первичной частицы генерируется 6000 событий. Было использовано пять полных выборок опорных ядер, протонов, гелия, азота, серы и железа.

Определение типа частицы. Каждому индивидуальному событию присваивается определенный тип частицы из промежутка ядер между легким протоном и тяжелым железом. Для смешанной выборки событий ШАЛ получается распределение этих событий по присвоенным массам ядер.

Оценка массы известных ядер. По полученному массовому распределению рассчитывается выборочное среднее значение массового числа для определенной энергии, диапазона зенитных углов, высоты расположения детектора. По выборочным средним происходит переход к истинным массам известных ядер. Данный этап необходим для минимизации ошибок оцененных масс.

Оценка средней массы по выборке. Для всей смешанной выборки по массовому распределению (полученному на втором этапе) рассчитывается выборочное среднее, по которому оценивается средняя масса.

8.1. Определение типа частицы

Задача о нахождении средней массы по смешанным выборкам событий ШАЛ решается методом «event-by-event». Каждому индивидуальному событию присваивается определенное значение традиционного критерия или логарифма массы, рассчитанный с помощью нейронной сети. Для всей смешанной выборки получается распределение логарифма массы по ядрам (рис. 9а). Плотности вероятности полученных распределений аппроксимируются Бета-распределением (1). Бета-распределение имеет два параметра **a**, **b** и определено в интервале от 0 до 1.

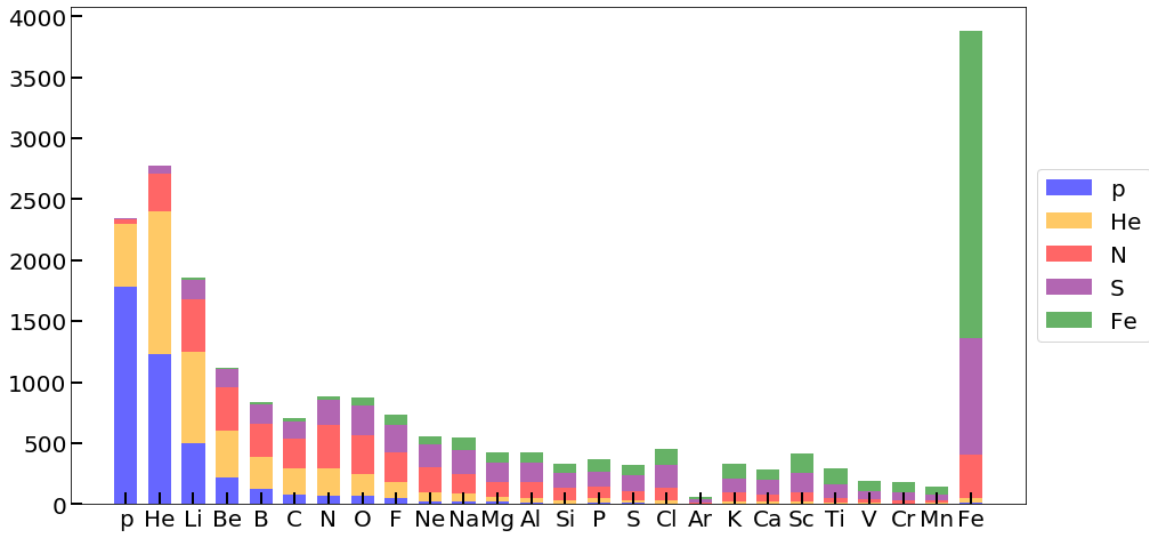


Рисунок 10. Массовое распределение

$$\begin{cases} F(\rho_{norm}) = \frac{\rho_{norm}^{a-1}(1 - \rho_{norm})^{b-1}}{\int_0^1 \rho_{norm}^{a-1}(1 - \rho_{norm})^{b-1} d\rho_{norm}} \\ \rho_{norm} \in [0, 1] \end{cases} \quad (1)$$

Всего между протоном и железом 26 ядер. Для каждого из ядер необходимо знать параметры Бета-распределения. По параметрам Бета-распределения (a, b , границы распределений) известных ядер проведены аппроксимации квадратичными функциями. Таким образом, найдены плотности вероятности для всех 26 распределений. На графике 9б) показаны плотности распределения известных ядер, рассчитанные по параметрам Бета-распределения, полученным в результате использования аппроксимаций. Отличия от оригинальных параметризованных распределений не превышают 10%.

На основе этого можно решить задачу определения типа частицы по максимуму плотности распределения. Каждому событию ШАЛ присваивается значение выхода массы нейронной сети. Будем выбирать одно из 26 ядер от протона до ядра железа. Для каждого ядра можем вычислить значение плотности вероятности для данного выхода массы. Выбираем ядро, имеющее наибольшую плотность вероятности. Таким образом, распределение делится на области, в которых значения выхода массы соответствуют определенному ядру. Следовательно, подсчитав значения выхода массы для любого события можно определить, какому типу частицы оно принадлежит.

В результате получается распределение смоделированных событий ШАЛ по массам. На гистограмме на рис. 10 цветом показано истинное ядро, а по оси абсцисс предсказанное.

8.2. Оценка массы известных ядер

По каждому из пяти известных массовых распределений (рис. 10) рассчитаны выборочные средние $\langle M \rangle$. На графике 11 отображена зависимость истинной массы A_{true} относительно полученного выборочного среднего. Разность между истинной массой и выборочным средним в зависимости от $\langle M \rangle$ аппроксимируется квадратичной функцией $fit(\langle M \rangle)$.

К каждому выборочному среднему добавляется значение в соответствии с этой аппроксимацией. Так получаем оценку массы известных ядер A_{est} .

$$A_{est} = \langle M \rangle + fit(\langle M \rangle) \quad (2)$$

В качестве оценивающей метрики использовано среднее абсолютное отклонение оцененных масс от истинных значений MAE . Оцененные массы в среднем отклоняются от реальных значений масс на 0.9 а.е.м. по пяти известным ядрам, рассчитанным по двум моделям адронного взаимодействия (рис. 12а).

$$MAE = \frac{\sum_{i=1}^5 |A_{est}^i - A_{true}^i|}{5}, \quad \text{где } i \text{ соответствует ядрам } p, He, N, S, Fe \quad (3)$$

С помощью универсального метода оценки массы по двум моделям адронного взаимодействия [14] для традиционного критерия получены оценки масс, которые отклоняются от истинных значений масс в среднем на 2.7 а.е.м. по пяти известным ядрам, рассчитанным по двум моделям ядерного взаимодействия (рис. 12б).

Из этого следует вывод, что у предложенного метода есть возможность не только оценивать среднюю массу по выборке, но и восстанавливать парциальные спектры космических лучей. Рассмотрим данный метод для расчета средней массы по всей выборке.

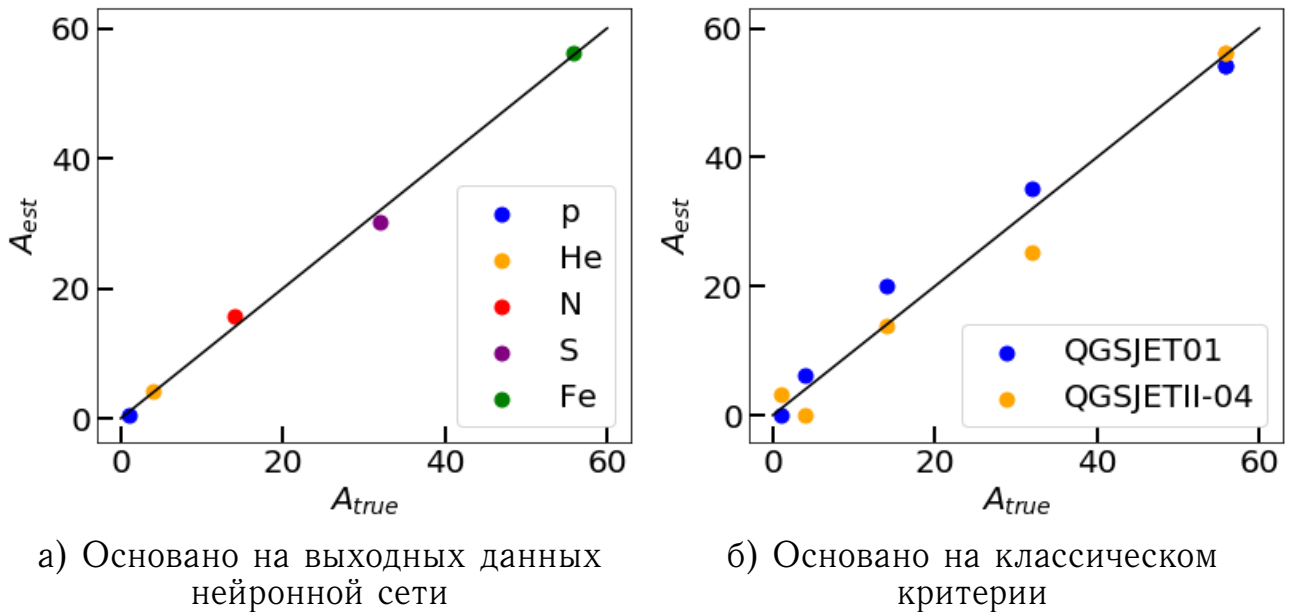


Рисунок 12. Оцененные массы A_{est} относительно истинных масс A_{true} ; сплошная линия – это диагональ под углом 45°

8.3. Точность восстановления средней массы по выборке из случайного количества разных ядер

С помощью описанного выше алгоритма найдены области распределения выхода массы, которые соответствуют максимуму распределения определенного ядра, и проведена оценка средней массы по смешанной выборке.

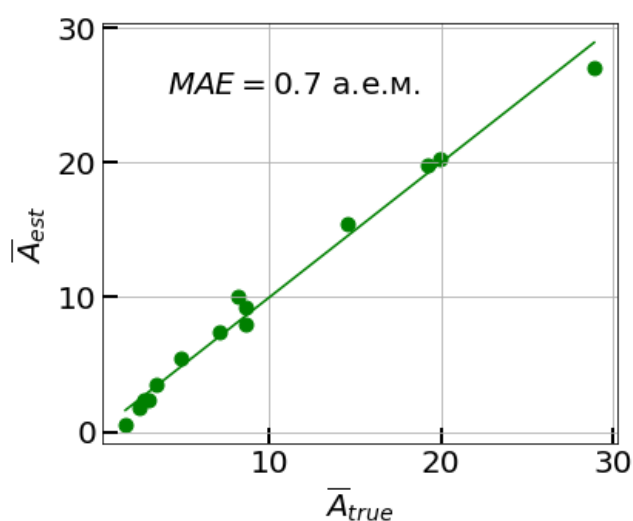
Составлена выборка событий по двум моделям адронного взаимодействия, в которой выбрано фиксированное количество случайных событий разных ядер, в том числе тех, которые не участвовали в поиске функций распределений выхода масс нейронной сети. Это ядра бериллия, углерода, натрия и кальция. Таким образом, зафиксировано истинное среднее массовое число.

По полученным массовым распределениям происходит расчет выборочного среднего \bar{M} по всей смешанной выборке событий с различными первичными ядрами. Прослеживается линейная зависимость рассчитанных выборочных средних от истинных средних масс разных выборок. Поэтому проводится аппроксимация зависимости сдвига истинных средних масс \bar{A}_{true} от выборочных средних масс по смешанным выборкам, в которых события разных ядер взяты в разных пропорциях. Полученная зависимость аппроксимируется линейной функцией $fit(\bar{M})$.

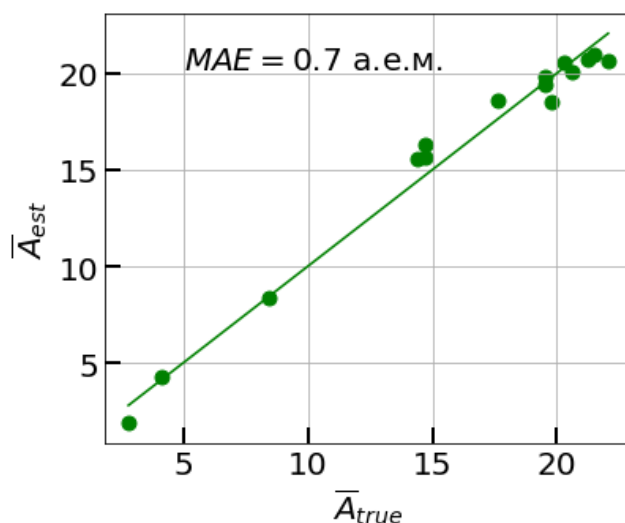
Из каждого выборочного среднего вычитается значение в соответствии с этой аппроксимацией. Так получаем оценку средней массы по смешанной выборке \bar{A}_{est} .

$$\bar{A}_{est} = \bar{M} - fit(\bar{M}) \quad (4)$$

На рис. 13а) приведена зависимость оцененных средних масс \bar{A}_{est} от истинных на основании выхода нейронной сети. Средняя абсолютная ошибка (MAE) оценки среднего массового числа по выборке равна 0.7 а.е.м. Средняя масса на основе классического критерия оценивается с идентичной ошибкой 0.7 а.е.м. Это означает, что с помощью детектора СФЕРА-2 среднюю массу более точно рассчитать не получится.



а) Основано на выходных данных нейронной сети



б) Основано на классическом критерии

Рисунок 13. Оцененные средние массы \bar{A}_{est} относительно истинных средних масс \bar{A}_{true} ; сплошная линия – это диагональ под углом 45°

Ошибка определения среднего массового числа меньше 1 а.е.м. является несравнимым результатом для экспериментов, нацеленных на изучение массовой компоненты ШАЛ в диапазоне энергий 1-100 ПэВ. В дальнейшем будет проведен эксперимент СФЕРА-3, более приспособленный для разделения первичных частиц по массам. Разработанный алгоритм получения оценок средних масс первичных частиц может быть успешно применен для экспериментов подобных СФЕРА.

§9. Парциальный спектр частиц

У такого метода, который приводит к лучшему разделению событий по массам есть возможность оценивать не только среднюю массу, но и проанализировать соотношение ядер, присутствующих в событиях.

В статье [25] решают такую задачу на основе данных нейтринной обсерватории IceCube. Метод обработки экспериментальных данных в таком случае следующий. В каждом срезе энергии гистограмма для каждого основного элемента (p, He, N, S, Fe) преобразуется в функцию плотности вероятности (PDF) и параметризуется (рис. 9б). Набор из пяти PDF затем взвешивается, чтобы найти дроби, которые лучше всего соответствуют выходной массе нейронной сети для экспериментальных данных в том же срезе по энергии. Таким образом, рассчитывается, как меняется содержание частиц от энергии.

§10. Выводы

В настоящей работе представлен эффективный метод разделения событий черенковского света атмосферных ливней по первичным массам с помощью нейронной сети на основе смоделированных данных аппарата СФЕРА-2 одновременно по двум моделям ядерного взаимодействия.

Разделение событий традиционным методом и методом глубокого обучения используют один и тот же набор моделирования. Было выбрано две модели ядро-ядерного взаимодействия, QGSJET01, QGSJETII-04. Выбрана атмосфера №1 в CORSIKA7 — аппроксимация Линсли стандартной американской атмосферы. Энергия первичной частицы составила 10 ПэВ, угол наклона оси ливня $10 - 20^\circ$, высота расположения детектора над уровнем снежной поверхности 900 метров.

В результате поиска лучшего разделения выхода нейронной сети архитектура AixNet была переделана под эксперимент СФЕРА-2. На вход нейронной сети подается видеоряд, в кадрах которого записаны количества фотоэлектронов в ФЭУ, двумерные карты нормализованных времен прихода ливня и суммарных сигналов атмосферного ливня. На выходе сети логарифм массы искомого ядра.

Классификация на 3 ядра (ядра водорода, азота и железа) методом глубокого обучения дает лучший результат, чем классификация классическими методами. На 20% лучше определяются ядра водорода и азота. Команда обсерватории Pierre Auger добилась результатов разделения протонов и ядер железа с ошибкой порядка 10% [23] для одной модели QGSJET. Применив нейронную сеть на смоделированных данных СФЕРА-2, ошибки разделения на два ядра составили порядка 4%. При этом модель обучалась одновременно на двух моделях ядерного взаимодействия. Таким образом, обеспечена независимость обработки экспериментальных данных от выбора модели ядерного взаимодействия.

Разделение событий по трем ядрам с помощью нейронной сети превосходит классический метод разделения более, чем на 10%. Благодаря чему становится возможным восстановление парциальных спектров космических лучей, в отличие от классического подхода на основе критериев.

Проведена оценка средней массы по смешанным выборкам. Средняя абсолютная ошибка определения средних масс по выборкам, в которых события различных ядер встречаются в разных пропорциях, равна 0.7 а.е.м. Средняя масса на основе классического критерия оценивается с идентичной ошибкой 0.7 а.е.м. Это означает, что с помощью детектора СФЕРА-2 среднюю массу более точно рассчитать не получится. Но даже с такой ошибкой задача о правильном восстановлении средней массы по спектру приписанных масс может быть успешно решена. Разработанный метод разделения событий атмосферных ливней по массам может быть применен к любому эксперименту типа СФЕРА и к наземным черенковским установкам.

Список литературы

1. *Abbasi R.U. et al.* // A Study of the Composition of Ultra-High-Energy Cosmic Rays Using the High-Resolution Fly's Eye // *Astrophys. J.* 2005. V. 622, N2. P. 910.
2. *Chernov D. V. et al.* // Primary Energy Spectrum and Mass Composition Determined with the Tunka EAS Cherenkov Array // *Mod. Phys. A.* 2005. V. 20, N29. P. 6799-6801.
3. *Tsunesada Yoshiki* // Telescope Array Composition Measurements // *Ibid.*
4. *Ivanov A.* // The Yakutsk Array Experiment: Main Results and Future Directions // *EPJ Web of Conf.* 2013. V. 53. P. 04003-04013
5. *Knurenko S. P.* // Cosmic Ray Spectrum in the Energy Range $1.0E15$ - $1.0E18$ eV and the Second Knee According to the Small Cherenkov Setup at the Yakutsk EAS Array // *Proc. of the 33rd ICRC, Rio de Janeiro, Brazil, July 2013.*
6. *Антонов Р.А. и др.* // Регистрация отражённого черенковского света ШАЛ в эксперименте СФЕРА как метод изучения космических лучей сверхвысоких энергий, // *Физика элементарных частиц и атомного ядра*, том 46, № 1, с. 115-166, 2015.
7. *Fowler J. W. et al.* // A Measurement of the Cosmic Ray Spectrum and Composition at the Knee // *Astropart. Phys.* 2001. V. 15, No. 1. P. 49-64.
8. *Chernov D. V. et al.* // Primary Energy Spectrum and Mass Composition Determined with the Tunka EAS Cherenkov Array // *Mod. Phys. A.* 2005. V. 20, No. 29. P. 6799-6801.
9. *Tokuno H. et al.* // The Cosmic Ray Primary Composition at the Knee Region from Lateral Distributions of Atmospheric Cherenkov Photons in Extensive Air Showers // *Astropart. Phys.* 2008. V. 29, No. 6. P. 453-460.
10. *Tsunesada Yoshiki* // Cosmic Ray Observation at Mount Chacaltaya for beyond the Knee Region // *Proc. of the 30th ICRC, Mexico City, 2008.* V. 4. P. 127-130.

11. *Apel W.D. et al.* // Ankle-Like Feature in the Energy Spectrum of Light Elements of Cosmic Rays Observed with KASCADE-Grande // *Phys.Rev.D*.2013.V.87. P.081101
12. *S. Thoudam*^{1,2}, *J. P. Rachen*¹, *A. van Vliet*¹, *A. Achterberg*¹, *S. Buitink*³, *H. Falcke*^{1,4,5} and *J. R. Hörandel*¹ et al. // Cosmic-ray energy spectrum and composition up to the ankle: the case for a second Galactic component, *A33*, V 595, 2016
13. *Чудаков А.Е.* // Возможный метод регистрации ШАЛ по черенковскому излучению, отраженному от заснеженной поверхности Земли // Экспериментальные методы исследования космических лучей сверхвысоких энергий: Материалы Всесоюз. симп., 19-23 июня 1972 г. Якутск. фил. Сиб. отд. АН СССР, 1974. С.69-74
14. *Латыпова В.С. Галкин В.И.* // Метод восстановления массового состава космических лучей высоких энергий (1-100 ПэВ) для детекторов типа СФЕРА // УЗФФ МГУ, 31.05.2023, статья в редакции, f56326
15. *Heck D.* et al. // CORSIKA Wissen-schaftliche Berichte. Forschungszentrum Karlsruhe 1998. FZKA 6019.
16. *N.N. Kalmykov, S.S. Ostapchenko and A.I. Pavlov* // *Nucl. Phys. B (Proc. Suppl.)* 52B (1997) 17
17. *S. Ostapchenko* // *Phys. Rev. D*89 (2014) 074009
18. *Tanguy Pierog, R. Engel, D. Heck* // Future of Monte Carlo Simulations of Atmospheric Showers // Karlsruhe Institute of Technology, AtmoHEAD 2014, Palazzo del Bo, Padova, Italy, 2014
19. *Erdmann M., Glombitza J., Walz D.* // A Deep Learning-Based Reconstruction of Air Showers at the Pierre Auger Observatory // *Astroparticle Physics*, V. 97, January 2018, P. 46-53.
20. *Guill'en A.* // Deep learning techniques applied to the physics of extensive air showers, // *Astroparticle Physics*, V 111, Pages 12-22, 2019
21. *Erdmann M., Glombitza J., Walz D.* // A deep learning-based reconstruction of cosmic ray-induced air showers // *Astroparticle Physics*, V 97, Pages 46-53, 2018
22. *Kingma D.P., Ba J.* // Adam: a method for stochastic optimization // *Proceedings of 3rd International Conference on Learning Representations (ICLR)*, San Diego, 2015. arXiv:1412.6980
23. *Kalashev O. et al* // Neural Networks as a Composition Diagnostic for Ultra-high Energy Cosmic Rays // *Journal of Instrumentation*, V 17, P05008, 2022
24. *M. Benedikt, A. Chance, B. Dalena, D. Denisov et. al.* // Future Circular Hadron Collider FCC-hh: Overview and Status // CERN, 16 march 2022.
25. *Aarsten M.G. et al. (IceCube Collaboration)* // Cosmic ray spectrum and composition from PeV to EeV using 3 years of data from IceTop and IceCube // *Phys.Rev.D* V 100 8, 082002, 2019