

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ

На правах рукописи

Голда Андрей Васильевич

**Методы машинного обучения
в исследовании электромагнитных
процессов рождения мезонов на протонах**

Специальность 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
Исупов Евгений Леонидович

Москва – 2026

Оглавление

Оглавление.....	2
Введение	5
1 Методы машинного обучения как инструмент для изучения физики ядра и частиц.....	13
1.1 Основные понятия машинного обучения	14
1.2 Ключевые аспекты обучения нейронных сетей.....	19
1.3 Методы предобработки данных и оценки качества работы алгоритма при обучении нейронных сетей.....	22
2 Исследование нейронных сетей в задаче моделирования дифференциальных сечений рассеяния и структурных функций для реакции электророждения одиночного пиона на протоне.....	25
2.1 Эксклюзивные реакции электророждения пиона на протоне. Дифференциальные сечения рассеяния и структурные функции.....	26
2.2 Модель машинного обучения для моделирования дифференциальных сечений в реакции $\gamma^* + p \rightarrow n + \pi^+$	31
2.2.1 Входные данные.....	31
2.2.2 Функция потерь	33
2.2.3 Отслеживаемые метрики в процессе обучения	34
2.2.4 Критерий сравнения точности моделирования с экспериментальной ошибкой измерения данных	35
2.2.5 Выбор модели.....	36
2.2.6 Архитектура нейронной сети	38
2.3 Процесс обучения нейронной сети	41
2.3.1 Аугментация данных	42
2.3.2 Стандартизация данных	43
2.3.3 Разделение набора данных на 3 модели	44

2.3.4	Введение весовых коэффициентов в функцию потерь	45
2.3.5	Условие ранней остановки обучения	46
2.3.6	Динамическая настройка скорости обучения	47
2.3.7	Версионирование экспериментов.....	47
2.3.8	Модификация функции потерь с учетом угловой зависимости дифференциальных сечений.....	48
2.3.9	Оптимизатор.....	49
2.4	Валидация результатов моделирования дифференциальных сечений и структурных функций	49
2.4.1	Сравнение результатов прогнозирования модели с экспериментальными данными CLAS.....	50
2.4.2	Сравнение результатов прогнозирования модели с феноменологической моделью MAID2007	57
2.4.3	Оценка статистической устойчивости модели методом реплик	61
3	Оценка сечений эксклюзивных каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$	67
3.1	Экспериментальные данные по электровозбуждению нуклонов	68
3.2	Методы оценки дифференциальных сечений электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах из экспериментальных данных CLAS	70
3.2.1	Формализм описания реакций электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах	71
3.3	Интерполяция эксклюзивных структурных функций	74
3.4	Экстраполяция структурных функций в каналах электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах	76
3.4.1	Экстраполяция структурных функций по W и Q^2	77
3.4.2	Экстраполяция структурных функций в область $3.45 < Q^2 < 5.00$ ГэВ ²	78
3.4.3	Экстраполяция структурных функций в области $Q^2 < 0.65$ ГэВ ²	80
3.5	Описание программы и веб-сайта для оценки сечений каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$	82

3.5.1	Комплекс программ для оценки сечений	84
4	Моделирование и анализ эксклюзивных каналов двухпионного фоторождения с применением алгоритмов машинного обучения.....	86
4.1	Двухпионное фоторождение.....	89
4.1.1	Описание реакции двухпионного фоторождения на протоне	89
4.1.2	Кинематика $\gamma p \rightarrow \pi^+ \pi^- p$	90
4.1.3	Двухпионное фоторождение на данных с детектора CLAS	93
4.2	Моделирование методом Монте-Карло	95
4.2.1	Генераторы событий фоторождения двух пионов	95
4.2.2	Моделирование детектора CLAS	100
4.3	Методология анфолдинга на основе GAN	105
4.3.1	Генеративно-состязательная сеть для моделирования детектора (DS-GAN)	107
4.3.2	Генеративно-состязательная сеть для анфолдинга детекторных эффектов (UNF-GAN)	109
4.3.3	Оценка неопределенности	111
4.4	Оценка качества моделирования процессов двухпионного фоторождения в эксклюзивных каналах	112
4.4.1	DS-GAN	113
4.4.2	UNF-GAN	117
	Заключение.....	121
	Список литературы.....	124

Введение

В настоящее время не вызывает никаких сомнений, что квантовая хромодинамика (КХД) является истинной теорией сильного взаимодействия. Эффективный инструмент для описания адронных процессов в КХД – теория возмущений – демонстрирует высокую предсказательную способность, что было многократно подтверждено экспериментально при изучении реакций с участием адронов в широком диапазоне энергий от десятков ГэВ до нескольких ТэВ и при величинах переданных импульсов от 10 до 10 000 ГэВ² [1].

Однако пертурбативные методы применимы только при высоких переданных импульсах или характерных расстояниях $r < 0.1$ Фм, характеризующихся значением безразмерного параметра $\alpha_s \ll 1$. С уменьшением переданных импульсов (или увеличением расстояния r) пертурбативный подход становится неприменимым, поскольку параметр α_s резко возрастает из-за эффекта антиэкранирования, обусловленного неабелевой природой КХД. При $\alpha_s \approx 1$ точечные кварки и глюоны формируют сложные структуры, включая виртуальные глюонные облака и кварк-антикварковые пары, что приводит к возникновению «одетых» частиц. Около 98% массы адронов формируется именно за счет процесса одевания, тогда как вклад механизма Хиггса не превышает 2% [2].

В промежуточной области ($0.1 \text{ Фм} < r < 1 \text{ Фм}$) особую роль играют возбужденные состояния нуклонов, изучение которых позволяет исследовать непертурбативную динамику сильного взаимодействия. Анализ спектра и структуры этих состояний, включая радиальные и орбитальные возбуждения, дает ценную информацию о формировании динамической массы и эволюции структуры одетых частиц [3].

Электромагнитное взаимодействие служит одним из основных инструментов исследования возбужденных состояний нуклонов. Значительная часть экспериментальных данных получена коллаборацией CLAS (Continuous Electron Beam Accelerator Facility - Large Acceptance Spectrometer) Национальной

лаборатории им. Т. Джефферсона в реакциях рассеяния электронов и фотонов на протонах и ядрах. Эти исследования направлены на изучение спектра резонансов в эксклюзивных реакциях фоторождения мезонов [4] и их внутренней структуры в процессах электророждения [5].

Получаемые данные включают дифференциальные сечения и поляризационные наблюдаемые, зависящие от множества кинематических переменных. Их анализ осложнен высокой размерностью фазового пространства, нелинейными корреляциями и статистическими флуктуациями. Традиционные статистические и параметрические методы не всегда позволяют полно описать структуру таких данных.

Методы машинного обучения дают возможность выявлять сложные многомерные зависимости, повышать точность анализа, что расширяет возможности исследования динамики сильных взаимодействий.

В последние годы Нобелевские премии по физике и химии были присуждены за выдающиеся достижения в области искусственного интеллекта, что демонстрирует, как методы ИИ становятся неотъемлемой частью современных научных исследований, открывая новые горизонты в понимании сложных систем и ускоряя прогресс в различных областях науки.

Таким образом, **актуальность темы** работы обусловлена необходимостью получения достоверной информации о наблюдаемых в электромагнитных процессах рождения мезонов на протонах в резонансной области и возможностью применения методов машинного обучения для моделирования многомерных экспериментальных данных, восстановления физических корреляций и учета детекторных эффектов.

Цели: Целью настоящей работы является создание методов анализа и обработки экспериментальных данных для выявления нелинейных физических корреляций в электромагнитных процессах с использованием алгоритмов машинного обучения на примере задачи многомерной интерполяции дифференциальных сечений и структурных функций в реакции электророждения

одионого мезона на протоне, а также задачи генеративного моделирования для восстановления детекторных эффектов в реакции фоторождения пар заряженных пионов.

Для достижения поставленных целей необходимо было решить следующие **задачи:**

1. Разработать метод вычисления значений структурных функций и дифференциальных сечений рассеяния реакции $\gamma^* + p \rightarrow n + \pi^+$ в резонансной области. Построить валидационную схему для определения качества работы предложенного метода.
2. Создать метод вычисления эксклюзивных структурных функций и дифференциальных сечений электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах. Оценить погрешности полученных результатов, а также области применимости данного метода. Провести сравнение оценок с экспериментальными данными.
3. Разработать метод, позволяющий устранять детекторные эффекты и выявлять корреляции в многомерном фазовом пространстве кинематических переменных для процессов фоторождения пар заряженных пионов на протонах. Провести процедуру тестирования метода на синтетических данных и определить систематические погрешности.

Основные положения, выносимые на защиту:

1. Применение нейросетевых подходов на основе полносвязной архитектуры с использованием независимых переменных, однозначно задающих кинематику реакции, реализует модельно-независимый метод оценки дифференциальных сечений реакции электророждения положительного пиона на протоне в резонансной области.
2. Разложение структурных функций в ряд по полиномам Лежандра, а также использование степенных функций для описания эволюции по квадрату четырехимпульса виртуального фотона Q^2 количественно определяет

зависимость дифференциальных сечений рассеяния и структурных функций от кинематических переменных в реакции электророждения $K^+\Lambda$ и $K^+\Sigma^0$.

3. Методы анализа и обработки экспериментальных данных, основанные на генеративно-состязательных нейронных сетях, выявляют корреляции в многомерном фазовом пространстве кинематических переменных и корректируют влияние эффектов, обусловленных откликом детектора, в событиях реакции двухпионного фоторождения.

Научная новизна:

1. Впервые разработан и реализован метод, основанный на полносвязных нейронных сетях, для моделирования экспериментальных данных по структурным функциям и дифференциальным сечениям рассеяния реакции $\gamma^* + p \rightarrow n + \pi^+$ в резонансной области, который воспроизводит данные с точностью, сопоставимой с экспериментальными погрешностями (средняя абсолютная ошибка 0.08 мкб/ср).
2. Создан метод вычисления эксклюзивных структурных функций и дифференциальных сечений электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах, позволяющий количественно определить зависимость этих величин от кинематических переменных.
3. Разработан метод выявления многомерных корреляций в событиях реакции $\gamma + p \rightarrow p + \pi^+ + \pi^-$ с использованием генеративных алгоритмов машинного обучения, позволяющий корректировать влияние детекторных эффектов и получать достоверные распределения четырехимпульсов исходных частиц.

Предмет и объект исследования. Объектом исследования являются электромагнитные процессы рождения мезонов на протонах, изучаемые в эксклюзивных реакциях электророждения и фоторождения мезонов. Предметом исследования являются методы машинного обучения для выявления нелинейных физических корреляций, моделирования дифференциальных сечений и структурных функций, а также коррекции детекторных эффектов в многомерном фазовом пространстве кинематических переменных.

Методология и методы исследования. Для моделирования дифференциальных сечений и структурных функций в реакциях электророждения мезонов на протонах использовались методы машинного обучения, включая полносвязные нейронные сети, методы многомерной регрессии и статистической валидации. Для анализа реакций электророждения $K^+\Lambda$ и $K^+\Sigma^0$ применялись разложения структурных функций по полиномам Лежандра и параметризация их зависимости от квадрата четырехимпульса виртуального фотона Q^2 . Для восстановления многомерных корреляций кинематических переменных и коррекции детекторных эффектов в реакции двухпионного фоторождения использовались генеративно-состязательные нейронные сети и методы Монте-Карло моделирования. Для численных расчетов в работе были реализованы алгоритмы на языках программирования C++ и Python. Язык Python также использовался для разработки и обучения нейронных сетей, проведения процедур их валидации, а также для анализа и визуализации полученных результатов.

Практическая значимость. Результаты работы моделей, полученные в диссертации, могут быть использованы для проверки точности теоретических моделей, а также для создания или улучшения работы существующих Монте-Карло генераторов событий. Также подобные алгоритмы могут послужить основой для системы хранения и распределения экспериментальных данных.

Достоверность результатов, описанных в диссертации, обеспечивается использованием апробированных и зарекомендовавших себя методов машинного обучения и классических методов обработки экспериментальных данных. Эффективность этих методов подтверждается согласием результатов с экспериментальными данными, полученными в ранее опубликованных работах коллаборации CLAS и мировых данных.

Апробация работы. Основные результаты работы докладывались автором на следующих конференциях: Международная молодежная научная конференция «Технологии ИИ в науке и образовании» (Москва - 2023), Научная конференция «Ломоносовские чтения» (Москва - 2024), Научная конференция «Ломоносов»

(Москва - 2024), 8-я международная конференция «Deep Learning in Computational Physics» (Москва - 2024), Научная конференция имени Б.С. Ишханова «Концентрированные потоки энергии в космической технике, электронике, экологии и медицине» (Москва - 2024).

Публикации. Основные результаты по теме диссертации изложены в 5 статьях, в том числе в 4 публикациях в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ имени М.В. Ломоносова по специальности и отрасли науки (через тире указан номер в списке литературы):

[A1] – [11] Alghamdi T., Alanazi Y., Battaglieri M., Bibrzycki L., Golda A.V., Hiller Blin A. N., Isupov E.L., Li Y., Marsicano L., Melnitchouk W., Mokeev V.I., Montana G., Pilloni A., Sato N., Szczepaniak. A. P., Vittorini T. Toward a generative modelling analysis of CLAS exclusive 2π photoproduction// Physical Review D. –2023. – Vol. 108, №9. – P. 094030. EDN: TVJNEE. Импакт-фактор 5,3 (JIF). 1,45 п.л.

[A2] – [37] Голда А.В., Голубенко А.А., Давыдов М.М., Исупов Е.Л., Мокеев В.И., Савкин С.А., Чистякова В.В. Оценка сечений эксклюзивных каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах из данных детектора CLAS. // Вестник Московского университета. Серия 3: Физика. Астрономия – 2024.– №4, – С. 2440201. EDN: UAXKEN. Импакт-фактор 0,651 (РИНЦ). 0,75 п.л.
Golda A.V., Golubenko A.A., Davydov M.M., Isupov E.L., Mokeev V.I., Savkin S.A., Chistyakova V.V. Cross section evaluation for exclusive channels of $K^+\Lambda$ and $K^+\Sigma^0$ electroproduction off protons using CLAS detector data. // Moscow University Physics Bulletin. –2024.– V. 79. № 4. – P. 450-461.

[A3] – [23] Golda A.V., Rusova A.A., Isupov E.L. et al. Machine Learning Approach in the Prediction of Differential Cross Sections and Structure Functions of Single Pion Electroproduction in the Resonance Region. // Moscow University Physics Bulletin. – 2024. –Vol. 79, №4. –P. S608-S621. EDN: LYGXAN Импакт-фактор 0,651 (РИНЦ). 0,95 п.л.

[A4] – [33] Голда А.В., Исупов Е.Л., Русова А.А., Чистякова В.В. Нейросетевое моделирование наблюдаемых реакций электророждения одиночного пиона на протоне в резонансной области // Вестник Московского Университета, – 2026. – №2. DOI: 10.55959/MSU0579-9392.81.2620204. Импакт-фактор 0,651 (РИНЦ). 1,04 п.л.

Прочие публикации:

[A5] – [25] Чистякова В.В., Голда А.В., Русова А.А., Исупов Е.Л. Машинное обучение в задаче предсказания дифференциальных сечений и структурных функций электророждения пионов в резонансной области // Ученые записки физического факультета Московского Университета – 2025. – №2. URL: <https://uzmu.phys.msu.ru/abstract/2025/2/2520213/>. Импакт-фактор 0,029 (РИНЦ). 0,42 п.л.

Личный вклад. Вклад соискателя во всех опубликованных работах является определяющим. Все выносимые на защиту положения являются результатами, полученными лично соискателем.

В публикации [A1] соискателем была реализована разработка и валидация нового подхода к восстановлению детекторных эффектов в эксклюзивной реакции двухпионного фоторождения с использованием генеративно-состязательных нейронных сетей. Соискателем были выполнены задачи анализа, связанные с построением и исследованием функций плотности вероятности для сгенерированных Монте-Карло псевдоданных, а также с проверкой применимости GAN-архитектур для раскрытия детекторных эффектов в многомерном фазовом пространстве. Проведенная соискателем работа являлась важной частью процедуры теста раскрытия и внесла вклад в обоснование основного вывода публикации о возможности применения генеративных моделей для восстановления детекторных эффектов при сохранении корреляций между кинематическими переменными.

В публикации [A2] соискателем был разработан и реализован метод вычисления эксклюзивных структурных функций и дифференциальных сечений

электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах из экспериментальных данных детектора CLAS, проведена интерполяция и экстраполяция сечений в широкой области кинематических переменных, выполнена оценка неопределенностей полученных результатов, а также реализован программный комплекс для расчета и визуализации дифференциальных сечений и структурных функций.

В публикации [A3] соискателем был предложен, реализован и валидирован передовой нейросетевой подход к моделированию дифференциальных сечений и структурных функций реакции электророждения одиночного пиона на протоне в резонансной области. Соискателем была разработана архитектура полносвязной нейронной сети, проведено обучение модели на экспериментальных данных CLAS, выполнена оптимизация гиперпараметров, а также проведено сравнение результатов моделирования с экспериментальными данными. В публикации [A4] соискателем была выполнена дополнительная валидация разработанного нейросетевого подхода, включая сравнение с феноменологической моделью MAID2007 и применение метода реплик для оценки устойчивости результатов и неопределенностей нейросетевого моделирования.

Соискателем была проведена основная работа при подготовке текста всех статей, а также представление статей в архив и редакции журналов и переписка с редакторами и рецензентами.

Структура и содержание работы. Диссертация состоит из введения, четырех глав, заключения и списка литературы. Полный объем диссертации составляет 135 страниц текста, включая 40 рисунков и 2 таблицы. Список литературы содержит 84 наименования.

1 Методы машинного обучения как инструмент для изучения физики ядра и частиц

С момента создания первых теоретических моделей искусственных нейронных сетей [6] до современных архитектур [7] машинное обучение (МО) существенно эволюционировало. Искусственный интеллект (ИИ) с течением времени стал неотъемлемым инструментом для исследователей в различных областях науки.

В последние годы Нобелевские премии по физике и химии были присуждены за выдающиеся достижения в области ИИ, что подчеркивает его растущее влияние на научные исследования. В 2024 году Нобелевская премия по физике была присуждена Джону Хопфилду и Джеффри Хинтону за их новаторскую работу над нейронными сетями, которые являются основой современных систем ИИ. Одним из таких алгоритмов является полносвязная рекуррентная нейронная сеть Хопфилда [8] – революционный алгоритм, способный восстанавливать закономерности из частичных или зашумленных входных данных, благодаря чему данный алгоритм сохраняет высокую эффективность даже при наличии неполных или зашумленных данных. Данный алгоритм, благодаря усовершенствованию Джеффри Хинтона, получил существенное развитие в стохастической рекуррентной архитектуре — машине Больцмана [9], где наличие скрытых внутренних представлений позволило эффективно решать широкий класс вычислительно сложных задач.

В области химии Нобелевская премия 2024 года была присуждена Дэвиду Бейкеру, Демису Хассабису и Джону Джамперу за их достижения в предсказании структуры белков с помощью ИИ. Их работа помогла раскрыть закономерности структуры белков, что имеет огромное значение в химии и инженерии. Особенно примечательна разработка модели искусственного интеллекта AlphaFold2 [10],

которая способна точно моделировать трехмерные структуры белков, что значительно ускоряет исследования в этой области.

Эти достижения демонстрируют, как методы искусственного интеллекта становятся неотъемлемой частью современных научных исследований, открывая новые горизонты в понимании сложных систем и ускоряя прогресс в различных областях науки.

Физика частиц не является исключением. Алгоритмы искусственного интеллекта используются для создания генераторов событий [11], подавления фоновых процессов, трекинга частиц [12], моделирования реакций, прогнозирования поведения динамических параметров и многого другого. Наличие больших наборов данных в физике частиц позволяет эффективно разрабатывать модели машинного обучения.

1.1 Основные понятия машинного обучения

Машинное обучение — это область искусственного интеллекта, которая разрабатывает алгоритмы, позволяющие выявлять закономерности в данных и делать прогнозы без явного программирования. Для разработки таких методов применяются инструменты математической статистики, численного анализа, оптимизации, теории вероятностей, теории графов, а также многие другие технологии обработки данных. МО называется «машинным», потому что компьютерная программа обучается автоматически, анализируя входные данные и корректируя свою модель на основе полученных результатов. Задачи решаются не напрямую, а посредством «обучения» на основе множества аналогичных примеров, когда модель итеративно изменяет внутренние параметры для повышения точности моделирования.

Основная идея машинного обучения заключается в создании алгоритмов, которые могут адаптироваться к новым данным. Вместо того чтобы писать строгие

правила, исследователи используют модели, которые учатся на основе объектов-примеров. Это особенно важно в физике частиц, где анализ больших массивов данных затруднен или невозможен при ручной обработке.

Объекты и признаки являются ключевыми понятиями машинного обучения:

- Объект — это отдельная независимая единица информации, на основе которой производится обучение модели, например, одна экспериментальная запись, реакция или событие.
- Признаки — это характеристики, описывающие объект. Для описания реакции такими признаками могут быть, например, количество частиц в конечном состоянии или кинематические переменные.
- Целевой признак — это та величина, которую модель машинного обучения должна моделировать в задачах классификации и регрессии (также называется «меткой» или «таргетом»). Когда говорят о размеченных данных, то в основном подразумевают наличие целевого признака, полученного из эксперимента или взятого на основе решения эксперта или из других источников.

Обычно в случае табличных данных каждый столбец (кроме целевого) — это признак, а каждая строка — объект. В контексте машинного обучения, а также в рамках настоящей работы, термины «предсказание», «прогнозирование» и «моделирование» используются как синонимы, обозначающие процесс получения оценок целевой переменной на основе входных признаков при помощи обученного алгоритма.

Машинное обучение включает в себя два основных типа:

- Обучение с учителем — это подход, при котором модель обучается на размеченных данных, где заранее известны правильные ответы. К таким задачам относятся классификация, когда нужно определить категорию объекта (например, определение, относится ли реакция к фоновым процессам или нет, — является типичным примером бинарной классификации), а также

регрессия, когда требуется спрогнозировать численное значение (например, угол вылета частицы на основе характеристик реакции).

- Обучение без учителя — это подход, который применяется, когда данные не имеют целевого признака, и модель самостоятельно ищет в них закономерности. Одним из примеров является кластеризация, при которой объекты разделяются на группы по сходным признакам, например, сегментация реакций по их типу. Также в данный тип обучения входит задача понижения размерности, целью которой является уменьшение числа признаков без потери информативности, что особенно полезно при работе со сложными многопараметрическими системами и наборами данных.

Также существуют и другие типы машинного обучения: обучение с подкреплением, частичное обучение с учителем, обучение с самообучением, а также алгоритмы, объединяющие эти подходы и другие типы.

Методы машинного обучения также обычно разделяют по видам алгоритмов на две большие группы: классическое машинное обучение и глубокое обучение.

- Классическое машинное обучение — это совокупность алгоритмов и подходов, разработанных до появления и широкого распространения глубокого обучения. Оно включает в себя алгоритмы с более интерпретируемым результатом, в основном нацеленные на работу с табличными данными. К ним относятся линейная и логистическая регрессия, деревья решений, байесовские классификаторы и многие другие.
- Глубокое обучение использует искусственные нейронные сети. Они требуют больших вычислительных ресурсов и объемов данных, но могут находить сложные зависимости, которые труднее выявить классическими методами, особенно при работе с неструктурированными данными (изображения, текст и другие).

Классические методы обычно легко интерпретируемы, а также требуют меньше вычислительных ресурсов по сравнению с нейросетями, поэтому остаются

актуальными для широкого спектра прикладных задач. Ниже представлены основные алгоритмы, относящиеся к классическому машинному обучению:

- Линейная регрессия — базовый метод, применяемый для моделирования числовых значений. Предполагает линейную зависимость между признаками и целевой переменной.
- Логистическая регрессия — используется для задач бинарной и многоклассовой классификации. Несмотря на название, относится к классификаторам и моделирует вероятность принадлежности объекта к классу.
- Метод ближайших соседей — основан на идее, что объекты с похожими признаками будут иметь схожие метки; работает без явного обучения, используя расстояние между объектами выборки.
- Классические методы кластеризации (например, K-means, DBSCAN) — используются для группировки данных без меток (обучение без учителя). Применяются для анализа структуры данных, выделения паттернов.
- Метод опорных векторов — эффективен для задач классификации и регрессии. Использует гиперплоскости, разделяющие классы с максимальным зазором, и может применять ядровые методы для работы с нелинейными разделениями.
- Деревья решений — это алгоритмы, основанные на рекурсивном разбиении пространства признаков на подмножества посредством последовательного применения правил, формируемых по критериям информативности. В процессе обучения строится иерархическая структура в виде дерева, в узлах которого находятся условия, позволяющие однозначно классифицировать или прогнозировать значение целевой переменной для каждого входного объекта.
- Ансамблевые методы — это подходы, объединяющие несколько моделей (обычно деревьев решений) для повышения точности. Включают механизмы бэггинга (например, случайный лес) и бустинга (например, XGBoost).

В отличие от классических алгоритмов, нейронные сети обладают разнообразными архитектурами, каждая из которых оптимизирована для определенных типов данных и специфических задач. Далее приведены основные виды нейросетевых архитектур, применяемых в современных системах машинного обучения:

- Полносвязные нейронные сети — архитектура сети, в которой каждый нейрон одного слоя связан с каждым нейроном последующего слоя. Исторически одна из первых архитектур, наиболее интерпретируемая и понятная.
- Сверточные нейронные сети — архитектура сети, предназначенная для обработки многомерных данных, в первую очередь изображений. Характеризуется применением сверточных фильтров, обеспечивающих выявление локальных признаков и инвариантность к сдвигам.
- Рекуррентные нейронные сети — нейросетевая архитектура, ориентированная на анализ последовательных данных за счет использования рекуррентных связей, позволяющих учитывать зависимость текущего состояния от предыдущих. Находит применение в обработке текстовых данных, прогнозировании временных рядов и машинном переводе.
- Автоэнкодеры — модель, обычно используемая для решения задач сжатия данных, обнаружения аномалий и снижения размерности. Базируется на симметричной структуре кодировщика и декодировщика, обучающихся восстанавливать входные данные.
- Трансформеры — архитектура, основанная на механизме самоорганизующегося внимания (self-attention). Применяется в решении широкого спектра задач обработки последовательностей, включая машинный перевод и генерацию текста.
- Генеративно-состязательные сети (Generative Adversarial Networks, GANs) — модель, включающая два взаимодействующих компонента (генератор и дискриминатор), обучение которых осуществляется в состязательной

парадигме. Может быть использована для генерации новых данных, синтеза изображений и улучшения качества обучающих выборок.

1.2 Ключевые аспекты обучения нейронных сетей

Обучение нейронных сетей представляет собой процесс оптимизации параметров модели с целью минимизации ошибки прогнозирования. Одним из фундаментальных элементов обучения является функция потерь, которая количественно оценивает расхождение между прогнозом модели и истинными значениями (значением целевого признака). В зависимости от типа и специфики задачи применяются различные функции потерь. Например, среднеквадратичная ошибка (Mean Squared Error, MSE) широко используется в задачах регрессии:

$$\mathcal{L}_{MSE} = \frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}, \quad (1.1)$$

где индекс i перебирает все значения из набора данных от 1 до N , y_i — фактическое значение целевой переменной, $\hat{y}_i$ — прогнозируемые значения целевой переменной. Тогда как функция перекрестной энтропии (Cross-Entropy Loss) чаще применяется в задачах классификации:

$$\mathcal{L}_{CE} = - \frac{\sum_j^N \sum_i^C y_i^j \log(\hat{y}_i^j)}{N}, \quad (1.2)$$

где индекс j перебирает все значения из набора данных от 1 до N , индекс i перебирает все классы (для случая задачи многоклассовой классификации) от 1 до C , y_i^j — фактическое значение принадлежности целевой переменной i -ого объекта к j -ому классу, $\hat{y}_i^j$ — прогнозируемое значение вероятности принадлежности i -ого объекта к j -ому классу. Функция потерь служит основным критерием, по которому модель оценивает качество моделирования и, соответственно, корректирует свои параметры.

Наиболее распространенным и эффективным методом коррекции параметров модели является метод градиентного спуска. Данный алгоритм является итеративным методом оптимизации и основан на вычислении градиента функции потерь по отношению к параметрам модели и на последующем обновлении этих параметров в направлении, уменьшающем значение функции потерь. Градиентный спуск может быть реализован в нескольких вариациях, включая стохастический градиентный спуск, мини-пакетный градиентный спуск и другие варианты, каждый из которых обладает своими преимуществами и недостатками в контексте скорости сходимости и вычислительной эффективности.

Основными параметрами, подлежащими оптимизации в процессе обучения, являются весовые коэффициенты и смещения. Веса представляют собой коэффициенты, регулирующие степень влияния входных данных на выходные значения нейросети, тогда как смещения позволяют модели адаптироваться к различным входным распределениям, обеспечивая дополнительную гибкость. Обновление весов и смещений является ключевым этапом процесса обучения, так как именно эти параметры определяют способность модели к обобщению входных данных. Абстрактно можно представить нейронную сеть в виде тензора чисел, где его форма, разреженность и размеры символизируют архитектуру сети. Такой тензор инициализируется случайным образом, а его элементы итеративно изменяются в процессе обучения сети. Теоретически, если эти параметры инициализируются «удачно», то необходимость обучения может быть устранена.

Процесс обучения выполняется итеративно, с многократным прохождением модели через весь обучающий набор данных. Полный проход по данным называется эпохой (epoch). В ходе каждой эпохи модель получает входные данные, выполняет прогнозирование, вычисляет функцию потерь, вычисляет градиент по этой функции и обновляет параметры. Количество эпох является критически важным гиперпараметром, поскольку недостаточное количество эпох может привести к недообучению (underfitting), а чрезмерное – к переобучению (overfitting), при котором модель теряет способность к обобщению и начинает запоминать обучающие объекты без выявления их закономерностей.

Для управления процессом подбора параметров используется оптимизатор, который определяет стратегию обновления весов. Среди наиболее популярных оптимизаторов можно выделить ADAM (Adaptive Moment Estimation), RMSprop (Root Mean Square Propagation) и SGD (Stochastic Gradient Descent). Выбор оптимизатора влияет на скорость сходимости модели и ее способность избегать локальных минимумов функции потерь. Оптимизаторы, основанные на адаптивных методах (такие как ADAM и RMSprop), демонстрируют высокую эффективность в задачах, связанных с большим объемом данных и сложной структурой входных признаков.

Помимо внутренних параметров модели, обучение управляется гиперпараметрами, которые задаются до начала процесса обучения и не изменяются в ходе оптимизации. К числу ключевых гиперпараметров может относиться, например, скорость обучения (величина шага, с которым модель обновляет свои параметры в направлении антиградиента функции потерь) или количество эпох. Выбор значений гиперпараметров оказывает существенное влияние на итоговую производительность модели, и их настройка часто осуществляется с помощью методов автоматизированного поиска, в том числе с использованием перебора по сетке, случайного поиска или байесовской оптимизации.

Таким образом, процесс обучения нейронных сетей представляет собой сложную итеративную процедуру, в которой задействованы методы оптимизации, статистические критерии оценки ошибки и алгоритмы управления параметрами модели. Оптимальная настройка данных компонентов обеспечивает точность прогнозирования и способность модели к обобщению, что является ключевым фактором для успешного применения таких моделей в различных областях машинного обучения. В наши дни не существует универсального алгоритма, который бы позволил подобрать оптимальный набор параметров вместе с архитектурой сети; таким образом процесс создания алгоритмов МО остается достаточно сложной исследовательской задачей.

1.3 Методы предобработки данных и оценки качества работы алгоритма при обучении нейронных сетей

Эффективное обучение нейронных сетей требует не только корректного выбора архитектуры и гиперпараметров, но и обеспечения высокого качества входных данных, включая их предварительную обработку. Методы оценки качества модели и предобработки данных существенно влияют на ее способность к обобщению и устойчивость к шуму. В данном разделе рассматриваются ключевые концепции, связанные с разделением данных, валидацией, обработкой выбросов, расширением обучающей выборки и стандартизацией.

Для успешного обучения модели изначальный набор данных разделяют на 3 части: тренировочную, валидационную и тестовую выборки. Тренировочным набором данных является часть исходного набора данных, которая используется для непосредственного обучения модели. На основе этого набора обновляются параметры модели в процессе обучения. Валидационная выборка представляет собой часть исходных данных, не используемую для обучения модели, но предназначенную для промежуточной оценки ее качества и подбора оптимальных гиперпараметров. Тестовая выборка предназначена для объективной оценки обобщающей работы алгоритма или качества работы модели после завершения обучения на ранее не использованных в обучении данных.

Одним из наиболее распространенных методов оценки обобщающей способности модели является перекрестная валидация. Данный метод используется для проверки качества модели на различных подмножествах данных, что позволяет снизить вероятность переобучения и дает более точные оценки. Наиболее популярной является k -кратная перекрестная валидация, при которой исходные данные делятся на k равных частей. Модель обучается на $(k - 1)$ частях и тестируется на оставшейся части. Процедура повторяется k раз, при этом каждый раз одна из частей данных используется в качестве валидационной, а остальные для обучения. Средний результат по всем итерациям используется как итоговая оценка

качества модели. Такой подход позволяет уменьшить зависимость от случайного разбиения данных и обеспечивает более стабильную и надежную оценку.

Одним из потенциальных источников ошибок в обучении является наличие выбросов – аномальных значений, которые значительно отклоняются от характерных значений в данных. Выбросы могут возникать по разным причинам, включая ошибки измерений, шум в данных, проблемы предобработки и иные источники, связанные, например, с некорректной разметкой или артефактами сбора данных. В зависимости от задачи выбросы могут быть удалены, скорректированы или учтены при обучении модели. Их присутствие способно существенно исказить распределение признаков, повлиять на значения функции потерь и, как следствие, привести к ухудшению обобщающей способности модели. В задачах регрессии выбросы могут приводить к смещению параметров модели, особенно при использовании методов, чувствительных к экстремальным значениям, таких как линейная регрессия с квадратичной функцией потерь. В задачах классификации наличие выбросов может усложнить границу принятия решений и вызвать переобучение.

Для идентификации выбросов применяются различные статистические методы и методы МО, включая правила трех сигм, межквартильный размах, методы кластеризации, локальные оценки плотности и алгоритмы выделения аномалий на основе деревьев. После выявления выбросов возможны разные стратегии: удаление выбросов, замена на статистические показатели (медиану, моду), введение дополнительных признаков, учитывающих их наличие, или использование устойчивых к выбросам алгоритмов. Выбор подхода зависит от характера данных, чувствительности модели к аномалиям и требований к точности.

Одной из важных стратегий улучшения качества обучения является расширение обучающей выборки (аугментация входных данных). Данный метод широко применяется, особенно в задачах компьютерного зрения, для искусственного увеличения объема обучающей выборки. Аугментация включает различные преобразования исходных данных, такие как повороты, отражения,

масштабирование, добавление шума или изменение цвета. Использование аугментации позволяет повысить устойчивость модели к изменениям входных данных, снизить вероятность переобучения и увеличить точность работы алгоритма.

Для обеспечения корректного представления данных перед подачей в модель используются методы стандартизации. Данный процесс заключается в преобразовании данных таким образом, чтобы их распределение имело среднее значение, близкое к нулю, и стандартное отклонение, близкое к единице. Стандартизация особенно важна при использовании градиентных методов оптимизации, так как значительные различия в масштабах признаков могут привести к замедлению сходимости модели или к ее неустойчивости.

Таким образом, методы кросс-валидации, корректное разбиение данных, обработка выбросов (и пропущенных значений), расширение обучающей выборки и стандартизация играют критически важную роль в повышении качества работы моделей машинного обучения. Их применение позволяет улучшить устойчивость методов МО к шуму, повысить их способность к обобщению и обеспечить стабильность прогнозирования в реальных условиях.

2 Исследование нейронных сетей в задаче моделирования дифференциальных сечений рассеяния и структурных функций для реакции электророждения одиночного пиона на протоне

В данной главе предлагается подход к обучению алгоритма на основе нейронной сети для прогнозирования значений дифференциальных сечений эксклюзивных реакций электророждения одиночного пиона на протоне в различных областях фазового пространства. Подобный алгоритм может служить основой для генераторов физических событий и представляет наибольшую ценность в режиме интерполяции и экстраполяции данных в ограниченной области фазового пространства.

Электророждение положительно заряженных пионов описывается реакцией:

$$\gamma^* + p \rightarrow n + \pi^+, \quad (2.1)$$

где источником виртуальных фотонов служит пучок электронов с энергией $E = 5.754$ ГэВ, $E = 5.499$ ГэВ, $E = 1.515$ ГэВ в зависимости от эксперимента.

Прогнозирование сечений как непрерывно меняющегося значения является задачей регрессии, и основной алгоритм, разработанный для ее решения, представляет собой глубокую полносвязную нейронную сеть с архитектурой из 19 скрытых слоев. Обучение проводилось с использованием функции потерь, которая не учитывает априорных теоретических знаний о процессе, что свидетельствует о модельной независимости предлагаемого алгоритма. Данные для обучения были взяты из экспериментов на детекторе CLAS национальной исследовательской

лаборатории имени Т. Джефферсона [13]. Они хранятся в базе данных CLAS Physics Database [14].

2.1 Эксклюзивные реакции электророждения пиона на протоне.

Дифференциальные сечения рассеяния и структурные функции

Первые исследования резонансов нуклона начались с изучения реакций упругого рассеяния π -мезонов на нуклонах [15]. Преимущество этого метода заключается в большом сечении взаимодействия мезона с нуклоном-мишенью. Однако налетающая частица (мезон) не является точечной, и ее внутренняя структура неизвестна. Более того, не все резонансные состояния могут быть обнаружены в упругом рассеянии π -мезонов, так как многие из них распадаются с эмиссией мезонов, отличных от π , или с эмиссией более одного мезона. Поэтому современные исследования структуры возбужденных состояний нуклонов проводятся в эксклюзивных реакциях, индуцируемых реальными и виртуальными фотонами [16].

Параметры резонансов нуклона включают его массу, ширину, квантовые числа, а также амплитуды фото- и электровозбуждения $A_{1/2}, A_{3/2}, S_{1/2}$. Вершина взаимодействия между виртуальным фотоном и протоном, ведущая к образованию резонанса, характеризуется тремя спиральными амплитудами $A_{1/2}, A_{3/2}, S_{1/2}$, которые зависят только от квадрата четырехимпульса, переданного фотоном. Индексы указывают на спиральность начальной системы частиц, то есть проекцию полного спина системы начальных частиц на направление импульса фотона в системе центра масс. Амплитуды $A_{1/2}, A_{3/2}$ соответствуют взаимодействию поперечно поляризованного фотона с нуклоном. Только эти амплитуды вносят вклад в фото-возбуждение резонансов нуклона. Амплитуда $S_{1/2}$ описывает

возбуждение резонансов продольно поляризованными виртуальными фотонами. Для реального фотона амплитуда $S_{1/2}$ равна нулю. Физические наблюдаемые, такие как сечения в системе центра масс, асимметрия фотонного пучка и асимметрия мишени, выражаются через сложные амплитуды реакции. Эти амплитуды представляют собой суперпозицию вклада резонансов нуклона и нерезонансных механизмов [17]. Диаграмма Фейнмана электромагнитного возбуждения нуклонных резонансов показана на рис. 2.1.

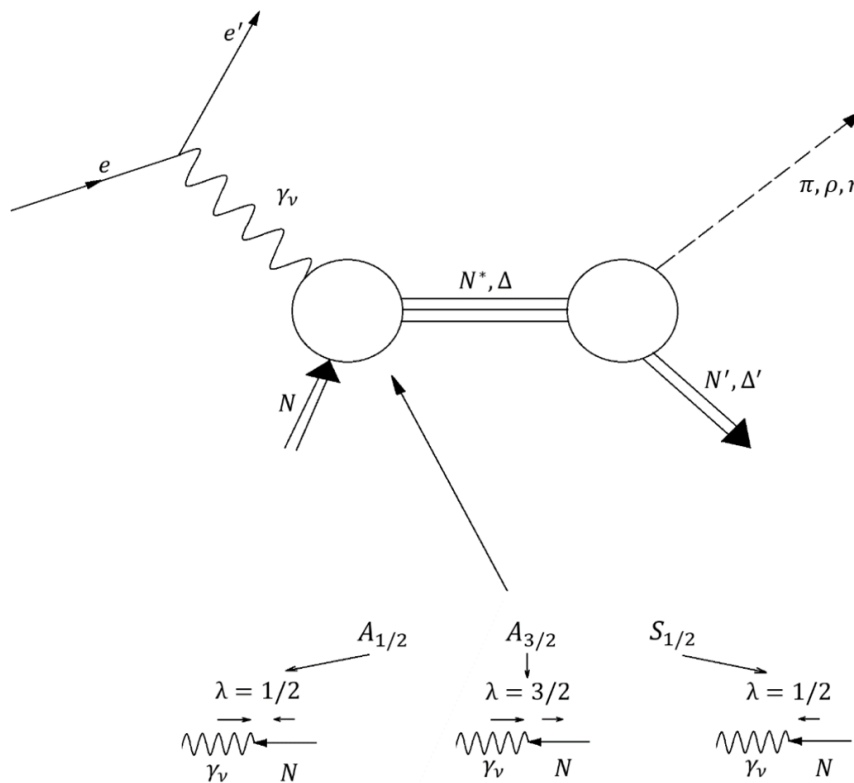


Рисунок 2.1 - Фейнмановская диаграмма электромагнитного возбуждения нуклонных резонансов.

В эксклюзивных реакциях измеряются кинематические характеристики всех частиц в конечном состоянии, и для удобства вводятся следующие переменные, характеризующие состояние системы: Q^2 — виртуальность фотона, W — инвариантная масса конечной адронной системы. Эксперименты по возбуждению нуклонов электронами позволяют изучать сечения при различных значениях виртуальности фотона Q^2 :

$$Q^2 = -q^2, \quad (2.2)$$

где q — переданный виртуальным фотоном четырехимпульс. Выражение для виртуальности фотона имеет вид:

$$\begin{aligned} Q^2 = -q^2 &= -(P - P')^2 = q^2 - (E - E')^2 = 2(EE' - |\vec{q}||\vec{q}'|\cos\theta) - 2m_e^2 \\ &\approx 2EE'(1 - \cos\theta_e) \approx 4EE'\sin^2\frac{\theta_e}{2}, \end{aligned} \quad (2.3)$$

где E — энергия электронного пучка, E' — конечная энергия электрона, а

$$\theta_e = \arccos \frac{p'_z}{E'} \quad (2.4)$$

— полярный угол рассеяния электрона в лабораторной системе координат. p'_z — проекция импульса рассеянного электрона на ось z . Во всех формулах предполагается, что скорость света равна единице:

$$c = 1. \quad (2.5)$$

Инвариантная масса конечной адронной системы W выражается через:

$$W^2 = (P_p + q)^2 = M^2 + 2M(E - E') - Q^2, \quad (2.6)$$

где M — масса протона, P_p — четырехимпульс протона.

Другими важными характеристиками рассматриваемой реакции (2.1) являются кинематические переменные θ_π^* и ϕ_π^* . Эта реакция рассматривается в системе центра масс фотона-протона. Схематическое изображение процесса представлено на рис. 2.2.

Угол θ_π^* — это полярный угол эмиссии пиона, а угол ϕ_π^* — угол между плоскостью реакции (плоскостью, образованной векторами импульсов мезона и нейтрона) и плоскостью рассеяния (плоскостью, образованной векторами импульсов падающего и рассеянного электрона).

Дифференциальные сечения рассеяния связаны с вероятностями переходов между начальными и конечными состояниями взаимодействующих частиц и измеряются экспериментально путем оценки скорости счета и распределения

импульсов частиц, образующихся в реакциях рассеяния на различных энергиях столкновения. Важно отметить, что в данной работе рассматриваются дифференциальные сечения рассеяния под воздействием виртуального фотона (2.1), а не сечения рассеяния электрона на протоне. Такой подход позволяет выделить структуру адронной системы и описывать ее независимо от параметров электронного пучка.

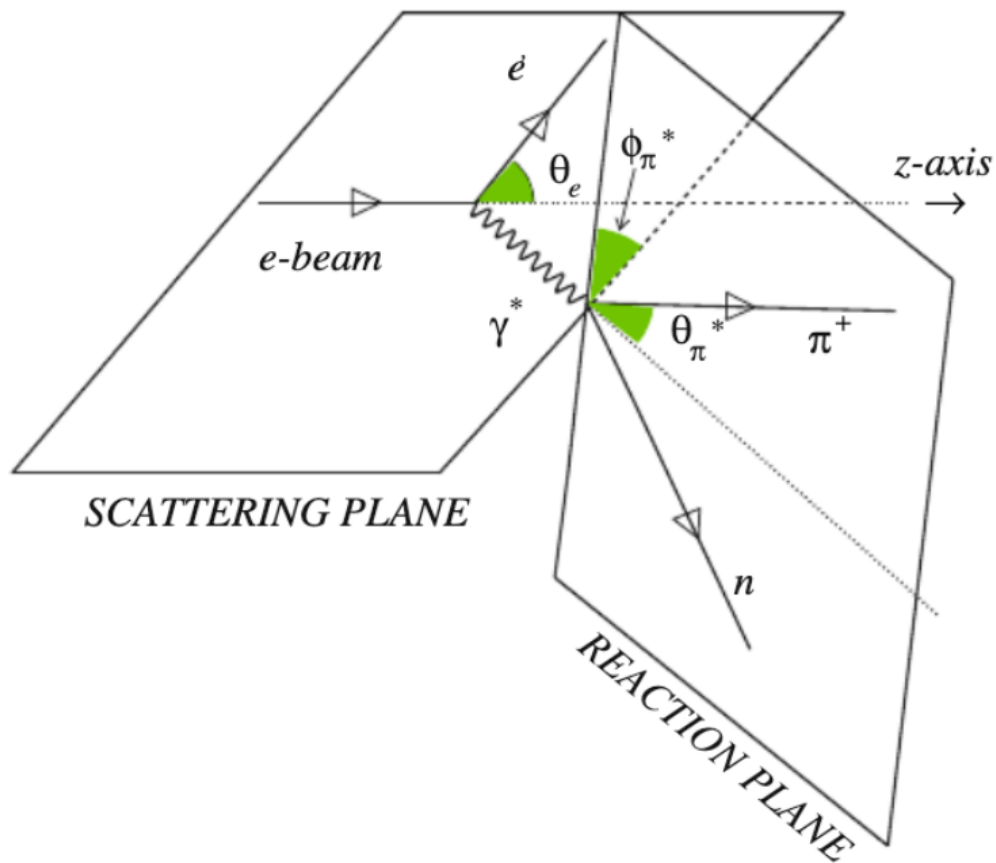


Рисунок 2.2 - Схематическое изображение кинематики реакции электророждения пиона на протоне.

В приближении однофотонного обмена, благодаря вращательной инвариантности амплитуд дифференциальные сечения рассеяния $\frac{d\sigma}{d\Omega}$ зависят от угла ϕ_π^* следующим образом [18]:

$$\begin{aligned}
& \frac{d\sigma}{d\Omega}(E, Q^2, W, \cos\theta_\pi^*, \phi_\pi^*) = \\
& = \frac{d\sigma_T}{d\Omega}(Q^2, W, \cos\theta_\pi^*) + \varepsilon \frac{d\sigma_L}{d\Omega}(Q^2, W, \cos\theta_\pi^*) + \varepsilon \frac{d\sigma_{TT}}{d\Omega}(Q^2, W, \cos\theta_\pi^*) \cos 2\phi_\pi^* + \\
& \quad + \sqrt{\varepsilon(1+\varepsilon)} \frac{d\sigma_{LT}}{d\Omega}(Q^2, W, \cos\theta_\pi^*) \cos\phi_\pi^* = \\
& = A(E, Q^2, W, \cos\theta_\pi^*) + B(E, Q^2, W, \cos\theta_\pi^*) \cos 2\phi_\pi^* + C(E, Q^2, W, \cos\theta_\pi^*) \cos\phi_\pi^*, \quad (2.7)
\end{aligned}$$

где ε — поперечная поляризация виртуального фотона, определяемая кинематикой рассеяния электронов на протонах, $A(E, Q^2, W, \cos\theta_\pi^*)$, $B(E, Q^2, W, \cos\theta_\pi^*)$, $C(E, Q^2, W, \cos\theta_\pi^*)$ — структурные функции

$$A(E, Q^2, W, \cos\theta_\pi^*) = \frac{d\sigma_T}{d\Omega}(Q^2, W, \cos\theta_\pi^*) + \varepsilon \frac{d\sigma_L}{d\Omega}(Q^2, W, \cos\theta_\pi^*),$$

$$B(E, Q^2, W, \cos\theta_\pi^*) = \varepsilon \frac{d\sigma_{TT}}{d\Omega}(Q^2, W, \cos\theta_\pi^*),$$

$$C(E, Q^2, W, \cos\theta_\pi^*) = \sqrt{\varepsilon(1+\varepsilon)} \frac{d\sigma_{LT}}{d\Omega}(Q^2, W, \cos\theta_\pi^*). \quad (2.8)$$

Структурные функции содержат основную информацию о структуре и динамике взаимодействия. Структурные функции, как показано в (2.7), определяются через интегрирование дифференциальных сечений рассеяния. Интегралы для вычисления структурных функций имеют следующий вид:

$$\begin{aligned}
A(E, Q^2, W, \cos\theta_\pi^*) &= \frac{1}{2\pi} \int_0^{2\pi} \frac{d\sigma(\phi_\pi^*)}{d\Omega} d\phi_\pi^*, \\
B(E, Q^2, W, \cos\theta_\pi^*) &= \frac{1}{\pi} \int_0^{2\pi} \frac{d\sigma(\phi_\pi^*)}{d\Omega} \cos 2\phi_\pi^* d\phi_\pi^*, \\
C(E, Q^2, W, \cos\theta_\pi^*) &= \frac{1}{\pi} \int_0^{2\pi} \frac{d\sigma(\phi_\pi^*)}{d\Omega} \cos\phi_\pi^* d\phi_\pi^* \quad (2.9)
\end{aligned}$$

Эти интегралы могут быть вычислены методом трапеций. После обучения алгоритма можно создать равномерную сетку параметров, одним из которых является угол ϕ_π^* . Основываясь на этой сетке, при помощи обученного алгоритма прогнозируются значения дифференциальных сечений рассеяния $\frac{d\sigma}{d\Omega}$. В случае

численного интегрирования методом трапеций эти интегралы заменяются на сумму, а функция $\frac{d\sigma}{d\Omega}$ аппроксимируется массивом значений, полученных из модели.

Исследование процессов электророждения одного мезона позволяет понять динамику процессов, происходящих в нуклонных резонансах. Низколежащие резонансы (с массами до 1.6 ГэВ) преимущественно распадаются по πN -каналу за счет сильного взаимодействия. Исследование однопионного канала необходимо для независимого изучения электромагнитных параметров в нуклонных резонансах [19, 20]. Диаграмма Фейнмана возбуждения резонансов нуклона через электромагнитный канал представлена на рис. 2.1.

Предметной областью исследования в данной работе является изучение непертурбативной области (инвариантная масса конечной системы адронов от 1 до 2 ГэВ), которая отвечает за формирование структуры нуклона и нуклонных резонансов.

2.2 Модель машинного обучения для моделирования

дифференциальных сечений в реакции $\gamma^* + p \rightarrow n + \pi^+$

2.2.1 Входные данные

Экспериментальные данные взяты из базы данных CLAS Physics Database [14]. Данные представлены в табличной форме, в дополнение к пяти независимым кинематическим переменным и самим дифференциальным сечениям рассеяния в наборе данных также указана экспериментальная ошибка определения дифференциальных сечений.

Набор данных содержит $\sim 10^5$ строк (98027 объектов), где каждая строка содержит 5 кинематических переменных (признаков): $E, W, Q^2, \theta_\pi^*, \phi_\pi^*$, одно

значение дифференциального сечения рассеяния (целевую переменную) $\frac{d\sigma}{d\Omega}$ и экспериментальную ошибку измерения дифференциального сечения $\varepsilon\left(\frac{d\sigma}{d\Omega}\right)$.

Хотя теоретически набор из этих пяти признаков полностью описывает кинематику реакции рождения положительно заряженного π -мезона на протоне при взаимодействии виртуального фотона, для моделирования дифференциальных сечений можно использовать любые входные данные, которые потенциально улучшают точность моделирования. Можно разрабатывать новые признаки на основе существующего набора данных, выводя новые переменные из физических принципов, или включать новые значения, такие как отклики детектора на аппаратном уровне, например, значение напряжения на выходном канале калориметра в момент регистрации частицы Ψ , или любые другие значения.

Однако в данной работе набор признаков был зафиксирован пятью указанными переменными, даже если это потенциально снижает качество модели за счет отказа от добавления новых признаков. Этот шаг был сделан по трем причинам:

1. На этапе прогнозирования нам необходимо предоставить модели тот же набор признаков, на котором модель была обучена, что может оказаться затруднительным или невозможным из-за трудностей с вычислением этих признаков. Например, если в обучающем наборе использовалось упомянутое выше значение напряжения на выходе калориметра в момент регистрации частицы Ψ , то для моделирования значений поперечного сечения конкретной точке фазового пространства нам также потребуется предоставить Ψ в качестве входных данных, даже если такие данные могут быть неизвестны до проведения реального эксперимента, что делает процесс прогнозирования невозможным.
2. Новые признаки, построенные на основе существующих пяти, такие как полиномиальные признаки (квадраты, кубы, четвертые степени) или

тригонометрические функции (синусы, косинусы), также, как и их суммы или взвешенные произведения, не вносят дополнительного вклада в точность обучаемых моделей. Это было подтверждено при проверке качества прогнозирования моделей с использованием кросс-валидации. В то же время поиск других более изощренных признаков проблематичен из-за того, что каждый новый признак требует валидации по соответствующим метрикам качества, что занимает много времени.

3. Использование только кинематических переменных реакции как признаков удобно для практического использования на этапе прогнозирования.

2.2.2 Функция потерь

На практике важно, чтобы алгоритм минимизировал разницу между прогнозируемыми и фактическими экспериментальными значениями. Если удастся достичь низкого значения этой разницы на валидационных и тестовых данных, то ожидается, что прогнозируемые значения также будут точны и в близкой кинематической области ($Q^2 < 5 \text{ ГэВ}^2$ и $W < 2.5 \text{ ГэВ}$), где отсутствуют экспериментальные данные. Этот достаточно простой и интуитивный принцип олицетворяет смысл функции потерь. Фактически, это именно та метрика, которую модель итеративно уменьшает в процессе обучения.

В данной работе используется классическая функция потерь – корень от среднеквадратичной ошибки (или метрика RMSE):

$$\mathcal{L}_{RMSE} = \sqrt{\frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}}, \quad (2.10)$$

где индекс i перебирает все значения из набора данных от 0 до N , y_i – экспериментальные значения целевой переменной (дифференциальных сечений), $\hat{y}_i$ – моделированные значения целевой переменной. Метрика RMSE выбрана по следующим причинам:

1. Стандарт в научных исследованиях и простота интерпретации. Метрика RMSE является общепринятой метрикой в работах, связанных с машинным обучением и физическим моделированием.
2. Пропорциональность квадратичному отклонению. Метрика RMSE минимизирует квадратичную ошибку между прогнозируемыми и истинными значениями, что особенно полезно, если необходимо придавать больший вес крупным ошибкам. Это позволяет модели фокусироваться на уменьшении значительных отклонений, что критично в задачах точных физических расчетов.
3. Метрика RMSE имеет хорошую математическую гладкость благодаря квадратам разностей, что делает процесс оптимизации более стабильным. Помимо этого, производная метрики RMSE имеет больший градиент при крупных ошибках (чем, например, градиент для функции MAE), что потенциально ускоряет обучение модели.

2.2.3 Отслеживаемые метрики в процессе обучения

В процессе поиска наиболее эффективной архитектуры нейронной сети и подбора гиперпараметров использовалась отложенная тестовая выборка, случайно отбираемая при каждом запуске из исходного набора данных. Размер тестовой выборки составлял 10% от общего количества входных данных - 8197 объектов.

Для оценки качества работы алгоритма использовалась метрика средней абсолютной ошибки (MAE):

$$M_{MAE} = \frac{\sum_i^N |y_i - \hat{y}_i|}{N}, \quad (2.11)$$

где индекс i перебирает все значения из набора данных от 0 до N , y_i - экспериментальные значения целевой переменной (дифференциальных сечений), $\hat{y}_i$ - моделированные значения целевой переменной.

Метрика MAE была выбрана по той причине, что это наиболее интерпретируемая метрика в задаче регрессии, которую просто использовать, с которой просто сравнивать качество моделей между собой и технически оценивать качество моделирования. Одновременно с простотой метрика MAE довольно точно подходит под постановку нашей задачи и под входные данные, которые практически не имеют выбросов и плотно расположены в пятимерном пространстве признаков. Помимо этого, метрика MAE также является стандартом в машинном обучении.

2.2.4 Критерий сравнения точности моделирования с экспериментальной ошибкой измерения данных

Входной набор данных содержит информацию о том, как оценить качество прогнозирования модели и ее применимость на практике. Это определяется величиной экспериментальной ошибки измерения дифференциальных сечений (целевой переменной) – $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$.

Стоит еще раз упомянуть, что на практике экспериментаторов интересует, чтобы алгоритм минимизировал разницу между прогнозируемыми и фактическими экспериментальными значениями. Если среднее отклонение прогнозируемых от фактических значений по средней абсолютной ошибке M_{MAE} оказывается меньше экспериментальной ошибки измерения дифференциальных сечений $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$, то можно считать, что алгоритм демонстрирует удовлетворительное качество моделирования и может считаться пригодным для практического применения.

2.2.5 Выбор модели

В данной работе задача заключается в прогнозировании вещественных чисел на основе заданного набора данных, что является задачей регрессии. Для ее решения классически прибегают к следующим существующим алгоритмам машинного обучения:

- линейные модели регрессии (базовый алгоритм линейной регрессии и его модификации с регуляризацией);
- метод опорных векторов;
- деревья решений;
- ансамблевые модели (случайный лес, XGBoost);
- подходы, основанные на полносвязных нейронных сетях.

В данной работе были протестированы все перечисленные алгоритмы на предоставленном наборе данных, и в качестве основного алгоритма был выбран нейросетевой подход по нескольким причинам.

Во-первых, классические алгоритмы машинного обучения, за исключением бустинга, не обеспечивают требуемого качества прогнозирования по критерию сравнения точности моделирования с экспериментальной ошибкой измерения данных, описанному ранее. Классические алгоритмы машинного обучения не смогли достичь качества ниже указанного порога по метрике средней абсолютной ошибки M_{MAE} .

Во-вторых, перечисленные выше алгоритмы не обладают достаточной гибкостью в процессе обучения. Физическая задача, поставленная в данной работе, включает множество вспомогательной информации, которую можно использовать для улучшения качества прогнозирования, однако не все алгоритмы могут ее учесть. Например, экспериментальная ошибка измерения дифференциальных сечений $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$ может определять степень “важности” каждого объекта в обучении

– чем ниже ошибка, тем точнее определены данные и тем более точного моделирования они требуют, тем самым вводя веса для объектов. Кроме того, известен вид зависимости дифференциальных сечений от угла ϕ_π^* (2.7), что также может быть использовано для улучшения качества прогнозирования.

В результате, для решения задачи у нас осталось два подхода, которые одновременно обладают достаточной точностью и гибкостью: основанный на нейронных сетях и основанный на бустинге. Алгоритм на основе бустинга был отвергнут по двум причинам:

1. Качество, полученное на модели XGBoost, оказалось ниже, чем на обученной полносвязной нейронной сети.
2. Отсутствие гладкости в моделировании. Даже после введения регуляризации, прогнозируемые значения казались переобученными, негладкими. Такое поведение может объясняться тем, что бустинг (особенно построенный на деревьях решений) — это не непрерывная модель. Модель фактически разбивает гиперплоскостями входное пространство на области и присваивает каждой фиксированное значение. Данный эффект можно наблюдать на рис. 2.3.

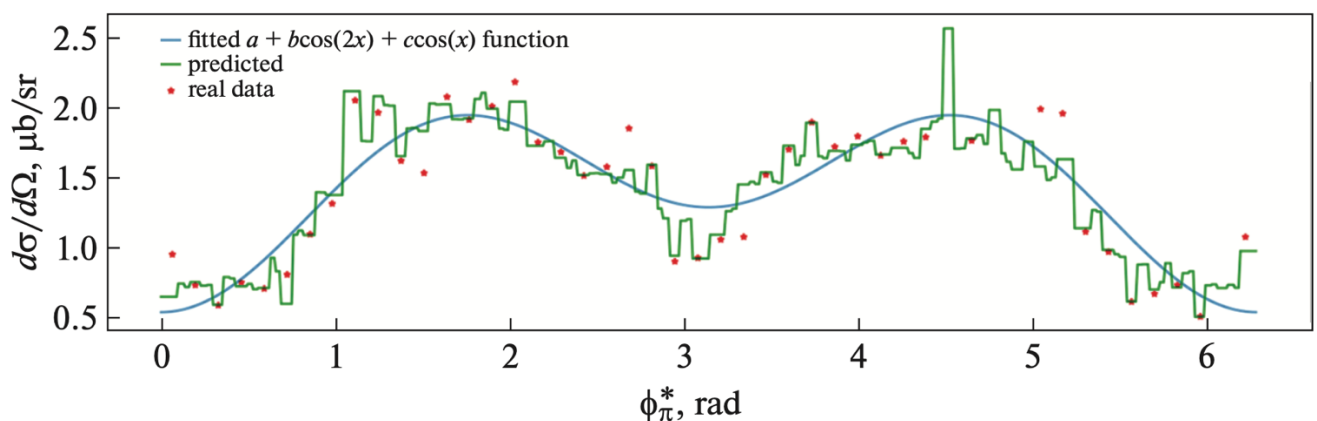


Рисунок 2.3 - Сравнение моделирования алгоритма на основе XGBoost (зеленая кривая) и экспериментальных данных для дифференциальных сечений рассеяния $\frac{d\sigma}{d\Omega}$ (красные точки) в зависимости от угла ϕ_π^* . Синяя кривая представляет собой фитирование экспериментальных данных согласно (2.7).

2.2.6 Архитектура нейронной сети

Существует широкий спектр нейронных сетей, работающих с табличными данными и решающих задачи регрессии [21]. В настоящее время нет универсального метода для определения того, насколько хорошо данная архитектура сети может прогнозировать значения целевой переменной. В то же время пространство возможных комбинаций архитектур и гиперпараметров неограниченно. В данной работе целью не ставилось создание архитектуры с самым точным результатом, наоборот, ставилась задача разработки достаточно простой архитектуры, способной решать задачу с точностью, не уступающей средней экспериментальной ошибке измерения дифференциальных сечений $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$.

Финальная модель представляет собой полносвязную нейронную сеть, состоящую из:

1. Входного слоя с 4 нейронами, соответствующими числу входных признаков (кинематических характеристик реакции): $W, Q^2, \theta_\pi^*, \phi_\pi^*$. Пятая кинематическая переменная E — начальная энергия электрона не использовалась напрямую как признак для отдельно взятой нейросети, но использовалась для разделения моделей по энергиям.
2. 19 скрытых слоев с переменным количеством нейронов, достигающим до 2000 в самом широком слое.
3. Выходного слоя, состоящего из одного нейрона, поскольку сеть решает задачу регрессии для дифференциального сечения реакции $\frac{d\sigma}{d\Omega}$.

Архитектура сети может быть представлена вектором (2.12), где каждый элемент определяет число нейронов в соответствующем по порядку слое. Также схематически архитектура сети представлена на рис. 2.4.

$$T = [4, 60, 80, 100, 120, 140, 240, 340, 440, 640, 2000, 1040, 640, 340, 240, 140, 100, 80, 60, 20, 1]. \quad (2.12)$$

Для функции активации использовалась стандартная ReLU (Rectified Linear Unit) [22]. Всего сеть содержит около $\sim 5 \times 10^6$ обучаемых параметров (весов и смещений).

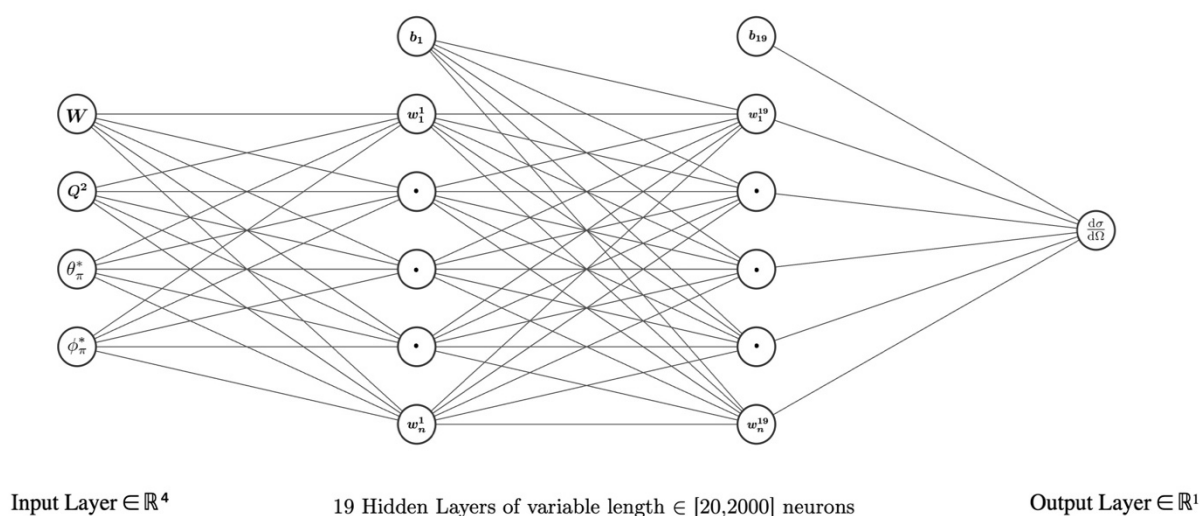


Рисунок 2.4 - Схематическое изображение архитектуры нейронной сети. Во входном слое (слева) находятся нейроны, соответствующие входным переменным, а в выходном слое (справа) – один выходной нейрон, соответствующий дифференциальному сечению рассеяния (целевой переменной). Скрытые слои имели переменную длину, b_i соответствуют смещению слоя, а w_j^i весам нейронов.

В данной работе не утверждается, что данная архитектура является оптимальной, однако хотелось бы объяснить принципы, которыми мы руководствовались при ее проектировании [23].

Во-первых, согласно теореме Цыбенко (теорема универсального аппроксиматора) [24], известно, что один скрытый слой в полносвязной нейронной сети может аппроксимировать любую нелинейную функцию. Однако такой единственный слой был бы очень большим (в терминах количества нейронов). Как

и во многих случаях, необходимо найти баланс между размером слоя и количеством слоев (глубиной сети).

Во-вторых, мы стремились ограничить время обучения одной сети, чтобы протестировать различные возможные архитектуры и режимы обучения с различными эвристиками. Поэтому изначально ставились ограничения примерно в 10^7 параметров — весов и смещений, что соответствует примерно двум часам обучения на современном оборудовании (практически любом современном ноутбуке) на представленном наборе данных со всеми описанными эвристиками.

На практике было выявлено, что точность результатов улучшается, если количество нейронов в сети монотонно увеличивается от входного слоя до центрального скрытого слоя (до 10-го слоя) и монотонно уменьшается от центрального скрытого слоя к выходному слою, создавая структуру сети в форме ромба. Также было установлено, что такая архитектура не оказывает негативного влияния на сходимость сети и время обучения.

Выбранная архитектура включает 19 скрытых слоев с переменным числом нейронов, что превышает глубину, обычно используемую в задачах регрессии на табличных данных, однако в данном случае такая глубина сети может быть оправдана сложной нелинейной зависимостью между входными параметрами и целевой величиной. Проведенный эмпирический анализ показал, что менее глубокие модели (с числом слоев от 5 до 10) немного уступают по точности по метрике средней абсолютной ошибки M_{MAE} , в то время как увеличение глубины сети до 19 слоев не приводит к переобучению — благодаря контролю валидационной ошибки. При этом архитектура сети остается достаточно простой: используются только полносвязные слои без усложняющих структур вроде слоев прореживания или пропускающих соединений, а общее число параметров сохраняется на уровне, приемлемом для эффективного обучения на доступном объеме данных (с учетом аугментации тренировочных данных).

Следует подчеркнуть, что целью данной работы не являлся поиск глобально оптимальной архитектуры нейронной сети. Основное внимание уделялось построению модели, способной обеспечить достаточно высокое качество моделирования, сопоставимое с уровнем экспериментальной погрешности, при относительно простой и воспроизводимой структуре. Вопрос выбора архитектуры, оптимальной одновременно по критериям точности, времени обучения и структурной компактности, требует отдельного систематического исследования. Проведение комплексного поиска по архитектурам (особенно в областях с небольшим количеством скрытых слоев) и гиперпараметрам, а также применение современных методов автоматизированного проектирования нейросетей, представляется перспективным направлением для последующих исследований [25].

Современные фреймворки глубокого обучения предоставляют широкую вариативность в реализации нейронных сетей. Разработка описанной нейронной сети в данной работе основана на фреймворке PyTorch [26].

2.3 Процесс обучения нейронной сети

Применение нейронных сетей данного типа не является новым; их основы были заложены в начале 1960-х годов вместе с основами метода обратного распространения ошибки [27]. Однако использование нейронных сетей в их классическом виде не дает удовлетворительных результатов. В ходе нашей работы было обучено и проверено множество вариантов сетей (около 200) с различными вариациями архитектуры и соответствующих гиперпараметров, включая различные эвристики. Ниже приведены наиболее эффективные эвристики, использованные в ходе обучения.

2.3.1 Аугментация данных

Аугментацией является процесс искусственного увеличения объема обучающей выборки путем создания новых данных на основе существующих, используемый для оптимизации точности прогнозирования алгоритма. Поскольку исходный набор данных содержит всего $\sim 10^5$ объектов (измерений дифференциальных сечений), то было решено искусственно увеличить его объем. Аугментация выполнялась методом выборки из нормального распределения. Для каждого объекта в наборе данных значения признаков немного сдвигались на случайную величину, взятую из гауссовского распределения с параметрами: μ — значение признака (уникальное для каждого отдельно взятого значения); σ — стандартное отклонение всех значений этого признака (константное для данного признака), деленное на коэффициент k_f . Данный коэффициент был выбран так, чтобы значения после аугментации не отклонялись значительно от оригинальных значений (менее чем на 1% от их величины), и в нашем случае $k_f = 30$.

Аугментация целевой переменной проводилась аналогично, но с другими параметрами гауссовского распределения. Так как известны значения экспериментальной ошибки измерения целевой переменной (дифференциальных сечений рассеяния), то значения целевой переменной при аугментации могут быть изменены в пределах всего диапазона ошибки. В нашем случае значения целевой переменной сдвигались на случайную величину из гауссовского распределения с параметрами: μ — значение целевой переменной (конкретного дифференциального сечения $\frac{d\sigma}{d\Omega}$); σ — экспериментальная ошибка измерения дифференциального сечения $\varepsilon\left(\frac{d\sigma}{d\Omega}\right)$ (определенная для конкретного дифференциального сечения $\frac{d\sigma}{d\Omega}$), деленная на коэффициент k_f . Данный коэффициент для целевой переменной был выбран $k_f = 3$, исходя из того принципа, что в пределах трех стандартных отклонений нормального распределения содержится $\sim 99\%$ данных. Таким образом,

в ~99% случаев при аугментации новые значения не выходили за пределы ошибки $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$.

Мы использовали коэффициент аугментации $m = 25$. Это означает, что при аугментации каждая строка набора данных была дублирована 25 раз с добавлением случайных значений из гауссовского распределения с соответствующими параметрами. Такое значение было выбрано как оптимальное с учетом компромисса между увеличением объема обучающей выборки и временем обучения модели.

2.3.2 Стандартизация данных

Стандартизация данных необходима для приведения всех признаков к единому масштабу, что способствует улучшению сходимости алгоритмов машинного обучения и повышает их точность. В частности, в нейросетях и других моделях, использующих градиентный спуск, данные с сильно различающимися диапазонами значений могут привести к неэффективному обучению: модель будет медленнее сходиться или вовсе не достигнет оптимума. Стандартизация помогает:

- Ускорить обучение модели. Данные с одинаковым масштабом позволяют алгоритму градиентного спуска двигаться более равномерно по всем направлениям, что делает процесс оптимизации более стабильным.
- Избежать доминирования одного признака над другими. Если признаки имеют разные порядки величин, например, один варьируется от 0 до 1, а другой от 10^3 до 10^6 , модель может опираться на большие числа, игнорируя маленькие, даже если те содержат более важную информацию.

В данной работе стандартизация проводилась путем вычитания среднего значения и деления на стандартное отклонение для каждого признака в отдельности.

2.3.3 Разделение набора данных на 3 модели

В начальном подходе обучалась единая нейросеть на всем наборе данных. Однако поскольку экспериментальные данные фактически представляют собой совокупность нескольких экспериментов, то данные формируют кластеры, сгруппированные по начальной энергии электронного пучка E , рис. 2.5. Эти кластеры имеют разные объемы данных, что приводит к «несбалансированному обучению»: сеть лучше распознает закономерности в одной области фазового пространства, но хуже в другой, а затем экстраполирует эти закономерности на весь диапазон, что приводит к снижению точности.

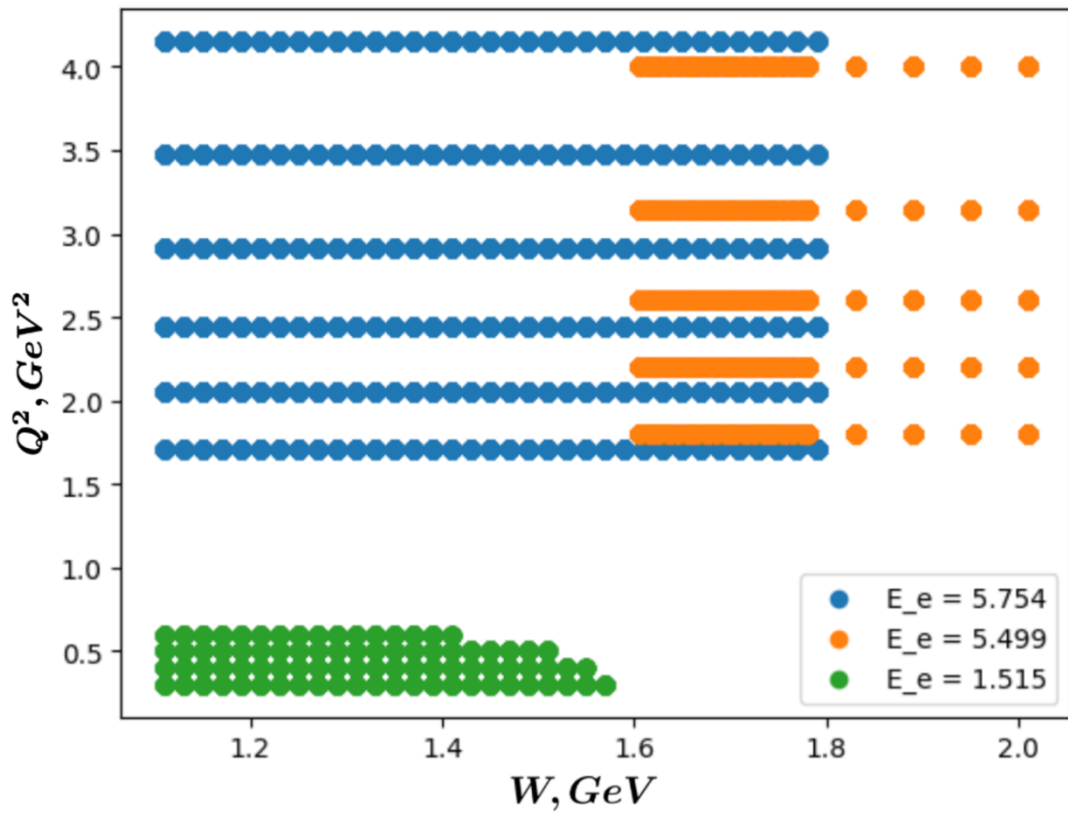


Рисунок 2.5 - График распределения экспериментальных данных в зависимости от виртуальности фотона Q^2 (ось ординат) и инвариантной массы конечной адронной системы W (ось абсцисс). Цветом указаны значения начальной энергии электрона E .

Зеленый для $E = 1.515$ ГэВ, оранжевый для $E = 5.499$ ГэВ, синий для $E = 5.754$ ГэВ.

Более того, при обучении единой сети наблюдался нежелательный эффект: в определенных диапазонах фазового пространства (особенно при W в интервале от 1.6 ГэВ до 1.8 ГэВ) прогнозируемые значения целевой переменной становились немонотонными. Это происходило из-за того, что при разных значениях энергии диапазоны W и Q^2 пересекались или лежали очень близко друг к другу, но значения целевой переменной (дифференциальных сечений) различались. Сеть улавливает эти несоответствия и переносит их на другие значения энергии, создавая артефакты в прогнозировании.

Поэтому было принято решение разделить исходный набор данных на три поднабора по значениям начальной энергии электрона ($E = 1.515$ ГэВ, $E = 5.499$ ГэВ, $E = 5.754$ ГэВ) и обучить три отдельные модели с одинаковой архитектурой сети и параметрами обучения. Такой подход значительно улучшил точность прогнозирования, как для отдельных моделей, так и в целом, устранив описанные выше проблемы.

2.3.4 Введение весовых коэффициентов в функцию потерь

Экспериментальная ошибка измерения дифференциальных сечений $\varepsilon \left(\frac{d\sigma}{d\Omega} \right)$ определяет вес (или значимость) каждого объекта в процессе обучения – чем ниже ошибка, тем точнее определены данные и тем выше требования к точности их моделирования. Объекты с меньшей ошибкой считаются более точными, и для них накладывается больший штраф за неправильные прогнозы.

Функция потерь RMSE $\mathcal{L}_{RMSE}$ имеет вид, указанный в (2.10). Ее модернизация - взвешенная функция потерь wRMSE определяется следующим образом:

$$\mathcal{L}_{wRMSE} = \sqrt{\frac{\sum_i^N w_i (y_i - \hat{y}_i)^2}{N \sum_i^N w_i}}, \quad (2.13)$$

где индекс i перебирает все значения из набора данных от 0 до N , y_i – экспериментальные значения целевой переменной (дифференциальных сечений), $\hat{y}_i$ – прогнозируемые значения целевой переменной, а w_i обратно пропорционально экспериментальной ошибке определения сечений $w_i = \frac{1}{\varepsilon_i\left(\frac{d\sigma}{d\Omega}\right)}$.

2.3.5 Условие ранней остановки обучения

Чтобы сократить время обучения и автоматизировать процесс, использовалась стратегия ранней остановки. Если значение функции потерь $\mathcal{L}_{RMSE}$ или $\mathcal{L}_{wRMSE}$ для валидационной выборки уменьшалось менее чем на 10^{-5} в течение последних 10 эпох, то процесс обучения завершался автоматически. Это позволяет избежать избыточного обучения, когда дальнейшее обучение не вносит значительного улучшения в качество модели. Пример кривых обучения для нейронной сети, которая обучалась на данных с начальной энергией электрона $E = 5.754$ ГэВ, представлен на рис. 2.6.



Рисунок 2.6 - Динамика функции потерь (RMSE) для обучающей и валидационной выборок при обучении нейронной сети для данных с энергией электрона $E = 5.754$ ГэВ.

2.3.6 Динамическая настройка скорости обучения

Скорость обучения — один из ключевых гиперпараметров, который необходимо тщательно настраивать для достижения оптимального результата. Начальная скорость обучения равна 0.001. Если средняя абсолютная ошибка M_{MAE} на валидационном наборе данных переставала снижаться (оставалась неизменной в течение последних 5 эпох), то скорость обучения уменьшалась в 2 раза. Такой метод обеспечивает более плавное и стабильное обучение модели. В нашем подходе использовалась динамическая корректировка скорости обучения, реализованная в библиотеке PyTorch.

2.3.7 Версионирование экспериментов

В процессе поиска наиболее подходящей конфигурации модели с разными архитектурами и комбинациями эвристик было обучено около 200 моделей. В такой ситуации возникла необходимость в системе хранения моделей и описания экспериментов. Для этого использовалась облачная платформа Weights & Biases [28], которая позволяет:

- хранить версии экспериментов,
- записывать параметры обучения и архитектуру сети,
- отслеживать все промежуточные и финальные метрики в структурированном формате (таблицы, графики).

Интерфейс платформы позволял быстро анализировать ключевые параметры обучения, включая значения функции потерь на тестовой выборке и среднюю абсолютную ошибку M_{MAE} , что способствовало отбору оптимальных конфигураций модели.

2.3.8 Модификация функции потерь с учетом угловой зависимости дифференциальных сечений

С физической точки зрения известно, что значения дифференциальных сечений зависят от угла ϕ_{π}^* в соответствии с уравнением (2.7). Так как модель прогнозирует дифференциальные сечения $\frac{d\sigma}{d\Omega}$, то потенциально можно вносить дополнительный штраф в функцию потерь за прогнозируемые значения, которые существенно отклоняются от этой зависимости.

Таким образом, функция потерь $\mathcal{L}_{RMSE}$ модифицируется следующим образом:

$$\mathcal{L}_{mRMSE} = \sqrt{\frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}} + \alpha \sum_i^N \frac{y_i - (A_i + B_i \cos(2\phi_i^*) + C_i \cos(\phi_i^*))}{N}, \quad (2.14)$$

где i — индекс, пробегающий все объекты выборки от 0 до N , y_i — значения целевой переменной, $\hat{y}_i$ — прогнозируемые значения целевой переменной, ϕ_i^* — угол между плоскостью реакции и плоскостью рассеяния ϕ_{π}^* для конкретного значения дифференциального сечения, A_i, B_i, C_i — значения структурных функций, полученные путем численного интегрирования зависимости дифференциальных сечений от угла ϕ_{π}^* (2.7) согласно (2.9) на интервале $[0, 2\pi]$, α — весовой коэффициент штрафного слагаемого, подбираемый эмпирически, принимал значение 0.2.

Хотя эта модификация функции потерь логична с физической точки зрения, финальные модели, представленные в данной работе, не включали эту эвристику. Это связано с тем, что включение штрафа не привело к значительному улучшению результатов прогнозирования, однако привязывало нашу работу к модельной зависимости (пусть и общеизвестной). Тем не менее, такой метод корректировки функции потерь хорошо изучен и широко применяется на практике [29, 30]. В дальнейшем такой подход потребует дополнительного тестирования и может быть реализован на следующих этапах проекта.

2.3.9 Оптимизатор

В качестве оптимизатора (алгоритм оптимизации, управляющий обновлением весов модели во время обучения) использовался алгоритм стохастической оптимизации ADAM [31], а конкретно его реализация на фреймворке PyTorch. Данный оптимизатор был выбран из-за достаточно высокой точности и скорости работы при обучении, что было особенно важно, поскольку в данной работе было обучено более 200 версий различных нейронных сетей.

2.4 Валидация результатов моделирования дифференциальных сечений и структурных функций

В данной работе исследуются методы искусственного интеллекта для прогнозирования дифференциальных сечений в эксклюзивных реакциях с рождением положительно заряженных пионов под воздействием виртуальных фотонов. На основе экспериментальных данных детектора CLAS был обучен набор полносвязных нейросетей без учета какого-либо теоретического знания о процессе рассеяния.

Был предложен набор методов валидации, включая метрику M_{MAE} , сравнение смоделированных дифференциальных сечений с экспериментальными данными в фиксированной области фазового пространства, а также сравнение смоделированных структурных функций с экспериментальными значениями в фиксированной области фазового пространства. Эти проверки демонстрируют, что нейросетевые алгоритмы обладают высокой точностью и позволяют интерполировать как дифференциальные сечения, так и структурные функции в различных областях фазового пространства. Для проверки качества прогнозирования проводится сравнение с феноменологической моделью MAID2007 [32], показавшее, что нейронная сеть корректно воспроизводит

характерную угловую зависимость и общее поведение сечений, включая области фазового пространства, не представленные в обучающей выборке. Также применяется метод реплик, основанный на построении ансамбля нейросетевых моделей, обученных на статистически флуктуирующих наборах данных, что позволяет количественно оценить устойчивость прогнозирования и связанные с ним неопределенности.

Подход на основе нейросетей сохраняет корреляции многомерного пространства кинематических переменных, является модельно-независимым и не использует априорных знаний о процессе. Он легко масштабируется на пространства высокой размерности, а также показывает высокое качество моделирования в широкой области фазового пространства, в том числе в областях с низкой плотностью экспериментальных данных. Ниже представлено обоснование того, что алгоритм способен предоставлять надежную интерполяцию данных. С прикладной точки зрения это делает данный алгоритм перспективной основой, например, для построения Монте-Карло генераторов событий или эффективных систем сбора и хранения данных.

2.4.1 Сравнение результатов прогнозирования модели с экспериментальными данными CLAS

Алгоритмы машинного обучения могут проверяться различными способами. Поскольку данная работа сосредоточена на конкретных физических процессах, а не на абстрактных задачах, в дополнение к классическим метрикам применяются физически мотивированные методы валидации.

Среди классических метрик качества модели была выделена одна основная - значение средней абсолютной ошибки (то есть среднего абсолютного отклонения прогнозируемых значений от реальных) M_{MAE} .

В результате проведения кросс-валидации, осуществленной на 10 разбиениях для каждой модели, при доле тестовой выборки 5% от общего объема исходных данных, средняя абсолютная ошибка моделирования составила $M_{MAE} = 0.08$ мкб/ср. Данное значение усреднено по трем моделям, каждая из которых обучалась на данных с разной начальной энергией электронного пучка, усреднение проводилось путем суммирования значений ошибки для всех разбиений и деления на 30. Для сравнения, среднее значение ошибки определения дифференциальных сечений в нашем наборе данных составляет $AVG\left(\varepsilon\left(\frac{d\sigma}{d\Omega}\right)\right) = \frac{\sum_i^N \varepsilon_i\left(\frac{d\sigma}{d\Omega}\right)}{N} = 0.2$ мкб/ср. Это свидетельствует о том, что моделированные значения дифференциальных сечений рассеяния в среднем находятся в пределах усредненного диапазона ошибки измерения самих экспериментальных данных.

В данном случае можно увидеть, что ошибка прогнозирования довольно мала, однако наиболее значимой с точки зрения физической интерпретации является проверка модели с использованием не только статистических основ, но и физических принципов. В качестве такой проверки строились распределения дифференциальных сечений рассеяния $\frac{d\sigma}{d\Omega}$ в зависимости от угла ϕ_π^* . На графиках (рис. 2.7, рис. 2.8) представлены экспериментальные данные (в ряде фиксированных областей фазового пространства) в виде красных точек. Синей кривой показан результат аппроксимации ϕ_π^* -зависимостью (2.7), где $A(E, Q^2, W, \cos\theta_\pi^*), B(E, Q^2, W, \cos\theta_\pi^*), C(E, Q^2, W, \cos\theta_\pi^*)$ являются параметрами фита. Зеленой кривой представлены значения прогнозирования модели нейронной сети. Черной кривой представлены значения прогнозирования сети, которые были использованы для восстановления сечений по ϕ_π^* -зависимости (2.7) следующим образом. На плотной сетке кинематических переменных выполнялось прогнозирование значений для целевой переменной $\frac{d\sigma}{d\Omega}$, на их основе высчитывались значения структурных функций $A(E, Q^2, W, \cos\theta_\pi^*), B(E, Q^2, W, \cos\theta_\pi^*), C(E, Q^2, W, \cos\theta_\pi^*)$ путем численного

интегрирования согласно (2.9), далее значения структурных функций подставлялись в выражение зависимости $\frac{d\sigma}{d\Omega}$ от ϕ_π^* (2.7), тем самым восстанавливая сечения $\frac{d\sigma}{d\Omega}$. Данный алгебраический прием был сделан для того, чтобы посмотреть, как ведут себя прогнозируемые значения целевой переменной, если алгоритму была напрямую указана ϕ_π^* -зависимость (2.7). Графики (рис. 2.7, рис. 2.8) демонстрируют, что прогнозирование модели (зеленая кривая) и прогнозируемые и восстановленные (черная кривая) дифференциальные сечения рассеяния достаточно близки к экспериментальным данным (красные точки) и фитированным данным (синяя кривая). Это показывает, что модель достаточно хорошо воспроизводит данные эксперимента.

В данной работе были проведены аналогичные сравнения для всего объема фазового пространства, где доступны экспериментальные данные (рис. 2.5). Во всем объеме фазового пространства моделирование демонстрировало высокую точность. Также следует отметить, что поскольку обученные алгоритмы хранят информацию из всего обучающего набора данных, в некоторых областях фазового пространства алгоритм обеспечивает более высокую точность интерполяции по сравнению с результатами фитирования. Указанный эффект можно наблюдать в областях фазового пространства с низкой плотностью экспериментальных данных, например, на втором графике снизу (рис. 2.7) и втором графике сверху (рис. 2.8). Представленные области соответствуют областям с низким или нулевым аксептансом детектора CLAS. Таким образом, построенная модель нейронной сети позволяет получать новую достоверную информацию о физической системе.

Кроме того, структурные функции могут быть однозначно получены из сечений реакции. На рис. 2.9 показано сравнение прогнозирования модели (черные точки) в зависимости от инвариантной массы конечной адронной системы с экспериментальными данными (красные точки) в различных областях фазового пространства. Такой метод может быть интерпретирован как сравнение контрольных сумм (относительно сравнений зависимости $\frac{d\sigma}{d\Omega}$ от ϕ_π^*), поскольку

структурные функции вычисляются численным интегрированием сечений. График (рис. 2.9) также демонстрирует, что прогнозируемые значения достаточно близки к экспериментальным данным. В данной работе также проводились аналогичные сравнения по структурным функциям для всего объема фазового пространства, где доступны экспериментальные данные (рис. 2.5). Высокая точность прогнозирования наблюдается по всему обследованному диапазону.

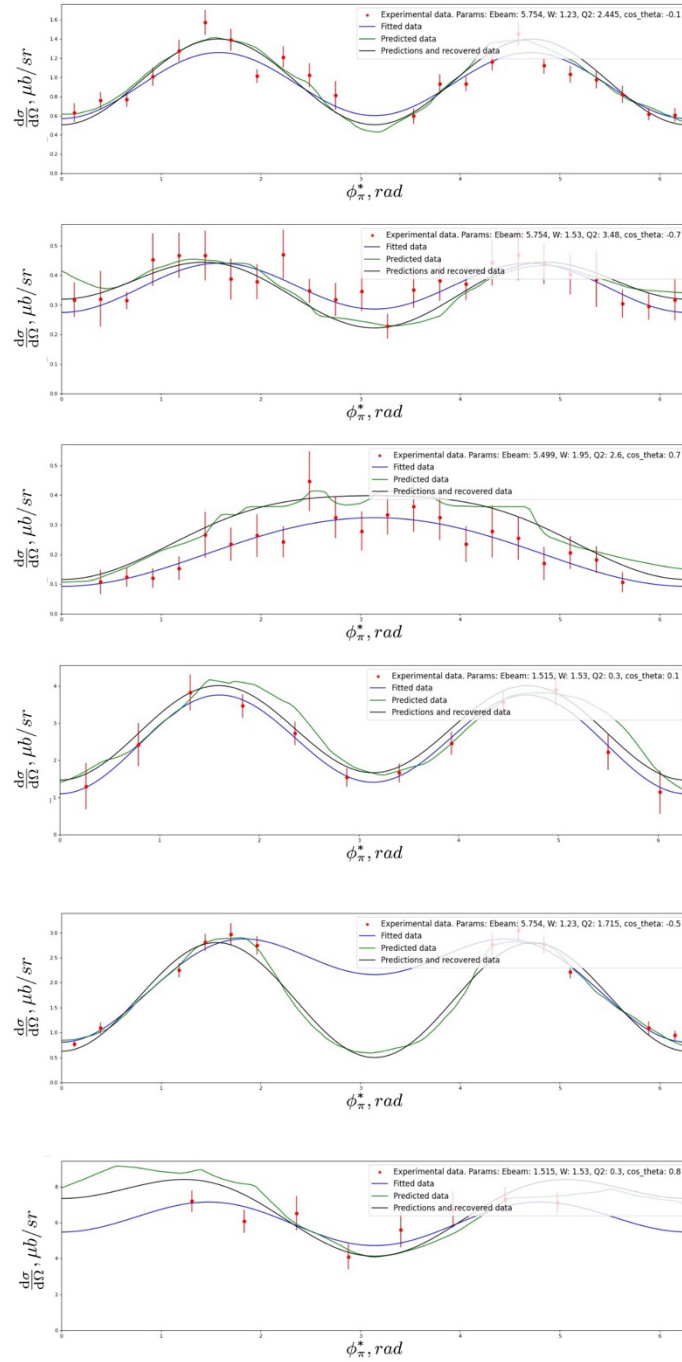


Рисунок 2.7 - Сравнение результатов работы модели для дифференциальных сечений рассеяния с экспериментальными данными в зависимости от угла ϕ_π^* в фиксированных областях фазового пространства, а именно:

$E = 5.754$ ГэВ, $W = 1.23$ ГэВ, $Q^2 = 2.445$ ГэВ², $\cos \theta_\pi^* = -0.1$ (первый график сверху),

$E = 5.754$ ГэВ, $W = 1.53$ ГэВ, $Q^2 = 3.48$ ГэВ², $\cos \theta_\pi^* = -0.7$ (второй график сверху),

$E = 5.499$ ГэВ, $W = 1.95$ ГэВ, $Q^2 = 2.6$ ГэВ², $\cos \theta_\pi^* = 0.7$ (третий график сверху),

$E = 1.515$ ГэВ, $W = 1.53$ ГэВ, $Q^2 = 0.3$ ГэВ², $\cos \theta_\pi^* = 0.1$ (четвертый график сверху),

$E = 5.754$ ГэВ, $W = 1.23$ ГэВ, $Q^2 = 1.715$ ГэВ², $\cos \theta_\pi^* = -0.5$ (пятый график сверху),

$E = 1.515$ ГэВ, $W = 1.53$ ГэВ, $Q^2 = 0.3$ ГэВ², $\cos \theta_\pi^* = 0.8$ (шестой график сверху).

Красными точками представлены экспериментальные данные CLAS. Зеленой кривой представлено прогнозирование нейронной сети. Синей кривой представлено фитирование данных согласно (2.7). Черной кривой представлены значения прогнозирования модели, восстановленные на основе выражения (2.7).

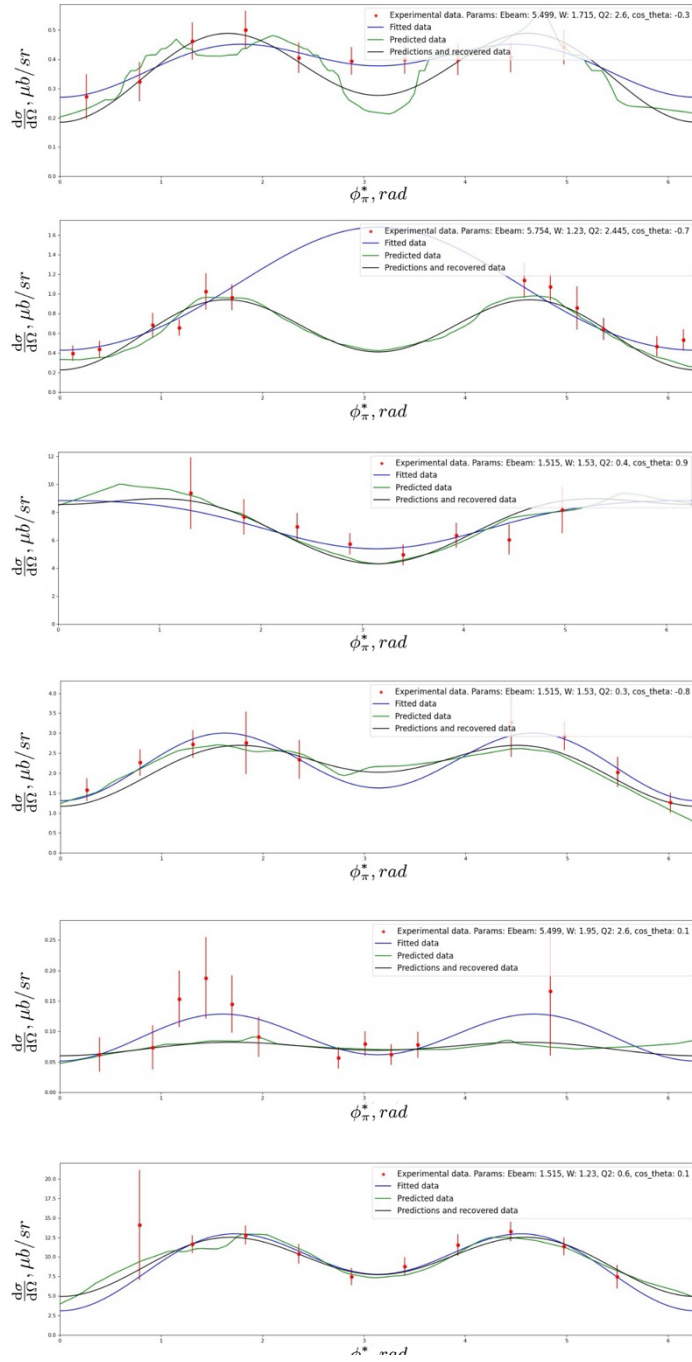


Рисунок 2.8 - Сравнение результатов работы модели для дифференциальных сечений рассеяния с экспериментальными данными в зависимости от угла ϕ_π^* в фиксированных областях фазового пространства, а именно:

$$E = 5.499 \text{ ГэВ}, W = 1.715 \text{ ГэВ}, Q^2 = 2.6 \text{ ГэВ}^2, \cos \theta_\pi^* = -0.3 \text{ (первый график сверху),}$$

$$E = 5.754 \text{ ГэВ}, W = 1.53 \text{ ГэВ}, Q^2 = 2.445 \text{ ГэВ}^2, \cos \theta_\pi^* = -0.7 \text{ (второй график сверху),}$$

$$E = 1.515 \text{ ГэВ}, W = 1.53 \text{ ГэВ}, Q^2 = 0.4 \text{ ГэВ}^2, \cos \theta_\pi^* = 0.9 \text{ (третий график сверху),}$$

$$E = 1.515 \text{ ГэВ}, W = 1.53 \text{ ГэВ}, Q^2 = 0.3 \text{ ГэВ}^2, \cos \theta_\pi^* = -0.8 \text{ (четвертый график сверху),}$$

$$E = 5.499 \text{ ГэВ}, W = 1.95 \text{ ГэВ}, Q^2 = 2.6 \text{ ГэВ}^2, \cos \theta_\pi^* = 0.1 \text{ (пятый график сверху),}$$

$$E = 1.515 \text{ ГэВ}, W = 1.95 \text{ ГэВ}, Q^2 = 0.5 \text{ ГэВ}^2, \cos \theta_\pi^* = 0.4 \text{ (шестой график сверху).}$$

Красными точками представлены экспериментальные данные CLAS. Зеленой кривой представлено прогнозирование нейронной сети. Синей кривой представлено фитирование данных согласно (2.7). Черной кривой представлены значения прогнозирования модели, восстановленные на основе выражения (2.7).

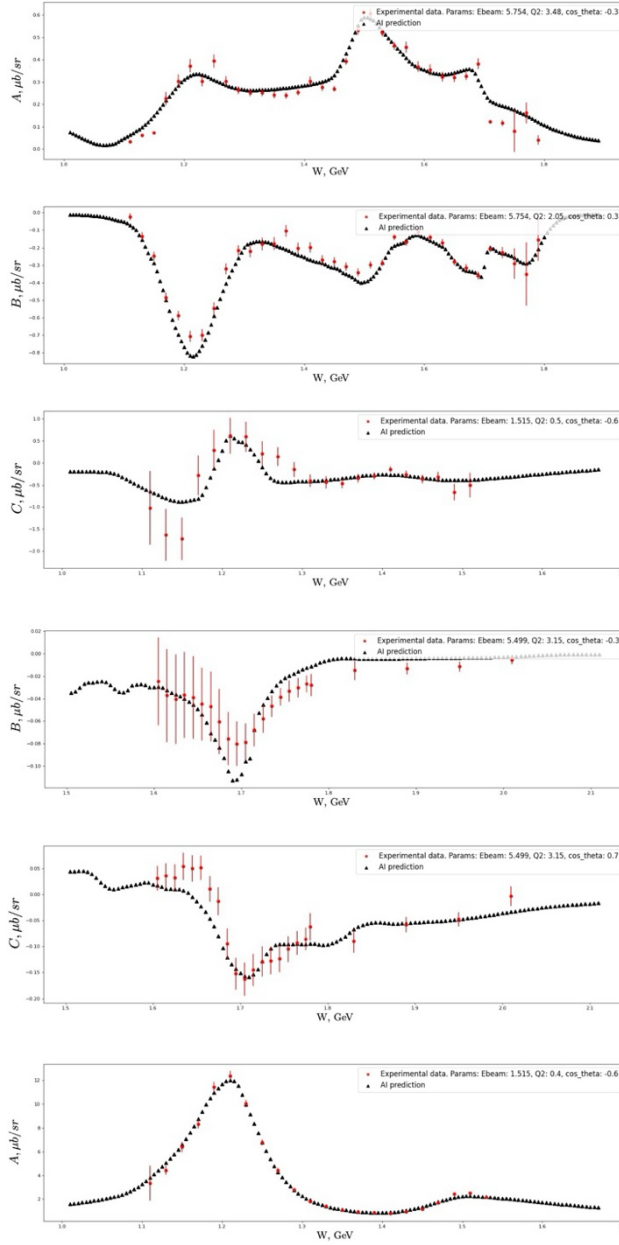


Рисунок 2.9 - Сравнение прогнозирования модели (черные точки) для структурных функций в зависимости от инвариантной массы конечных адронов W с экспериментальными данными (красные точки с ошибками) в фиксированных областях фазового пространства, а именно:

Функция A при $E = 5.754$ ГэВ, $Q^2 = 3.48$ ГэВ², $\cos \theta_\pi^* = -0.3$ (первый график сверху),

Функция B при $E = 5.754$ ГэВ, $Q^2 = 2.445$ ГэВ², $\cos \theta_\pi^* = -0.7$ (второй график сверху),

Функция C при $E = 1.515$ ГэВ, $Q^2 = 0.5$ ГэВ², $\cos \theta_\pi^* = -0.6$ (третий график сверху),

Функция B при $E = 5.499$ ГэВ, $Q^2 = 3.15$ ГэВ², $\cos \theta_\pi^* = -0.3$ (четвертый график сверху),

Функция C при $E = 5.499$ ГэВ, $Q^2 = 3.15$ ГэВ², $\cos \theta_\pi^* = 0.7$ (пятый график сверху),

Функция A при $E = 1.515$ ГэВ, $Q^2 = 0.4$ ГэВ², $\cos \theta_\pi^* = -0.6$ (шестой график сверху).

2.4.2 Сравнение результатов прогнозирования модели с феноменологической моделью MAID2007

Для дополнительной проверки физической корректности нейросетевого подхода было выполнено сравнение с феноменологической моделью MAID2007. В отличие от прямого сопоставления предсказаний нейронной сети и модели в произвольных точках фазового пространства, в настоящей работе использовалась более строгая схема валидации - нейронная сеть обучалась на значениях дифференциальных сечений, рассчитанных в рамках модели MAID2007 в тех же кинематических точках, где в исходной задаче располагались экспериментальные данные CLAS [33]. Такая постановка позволяет проверить, способна ли выбранная архитектура нейронной сети корректно восстановить многомерную функциональную зависимость, заданную известной феноменологической моделью, и не приводит ли сама процедура нейросетевой аппроксимации к нефизическим искажениям.

Модель MAID2007 представляет собой унитарную изобарную модель, широко используемую для описания процессов фото- и электророждения пионов на нуклонах в резонансной области. В ней учитываются как резонансные, так и нерезонансные вклады в амплитуду реакции. Поэтому сравнение с моделью MAID2007 удобно использовать как независимую проверку нейросетевого алгоритма, особенно в задаче моделирования рассматриваемой реакции (2.1) в резонансной области.

Схема сравнения была организована следующим образом. Сначала для набора кинематических точек, совпадающего с расположением экспериментальных измерений CLAS, с помощью модели MAID2007 были вычислены значения дифференциальных сечений. Далее именно эти значения использовались как целевая переменная при обучении нейронной сети. Тем самым исключалось влияние экспериментального статистического шума и проверялась способность сети воспроизводить известную гладкую зависимость, задаваемую

феноменологической моделью. После обучения результаты нейросетевой модели сравнивались с исходными значениями модели MAID2007 как в самих узлах обучающей выборки, так и в промежуточных областях фазового пространства, где явные обучающие точки отсутствовали. Такая процедура фактически позволяет проверить интерполяционные свойства модели и оценить, насколько корректно сеть восстанавливает кинематическую структуру сечений.

Следует отметить, что в рамках данной проверки использовалась та же архитектура нейронной сети, что и в основном анализе экспериментальных данных, за исключением лишь того, что здесь мы не проводили разделения сетей на 3 независимых модели по начальной энергии пучка электронов. В процессе обучения контролировалась динамика функции потерь на обучающей и валидационной выборках. Анализ кривых обучения (рис. 2.10) показывает устойчивую сходимость. Значение средней абсолютной ошибки (MAE) на тестовой выборке составляет порядка 0.01 мкб/ср. Более низкое значение средней абсолютной ошибки по сравнению с обучением на экспериментальных данных связано с тем, что феноменологическая модель MAID2007 формирует менее волатильные значения дифференциальных сечений без статистических флуктуаций и экспериментальных погрешностей, вследствие чего соответствующая зависимость легче аппроксимируется нейронной сетью.

Особый интерес представляет зависимость дифференциального сечения от азимутального угла ϕ_{π}^* (2.7). Как и в случае анализа экспериментальных данных, именно эта зависимость позволяет наиболее наглядно проверить, сохраняет ли нейросетевая модель правильную форму распределения.

Описанное сравнение было выполнено во всей области фазового пространства, где есть экспериментальные данные, включая участки, где между соседними экспериментальными точками расстояние по кинематическим переменным относительно велико (в частности, для переменной ϕ_{π}^*). В этих условиях особенно важно проверить, не возникает ли у нейросетевой модели неустойчивого поведения между обучающими точками. Полученные результаты

показали, что согласие между нейросетевой моделью и MAID2007 сохраняется не только в самих узлах обучающей выборки, но и в промежуточных областях. Иными словами, сеть воспроизводит не просто дискретный набор значений, а именно многомерную функциональную зависимость, заложенную в феноменологической модели. Данное наблюдение является принципиально важным, поскольку одна из основных задач разработанного алгоритма состоит именно в интерполяции сечений в непрерывном фазовом пространстве. На рис. 2.11 показан пример сравнения в области фазового пространства, где акцент сделан именно на поведении модели между узлами исходной сетки. Видно, что нейросетевое предсказание остается гладким и не содержит нефизических осцилляций, а его поведение согласуется с формой зависимости, задаваемой моделью MAID2007.

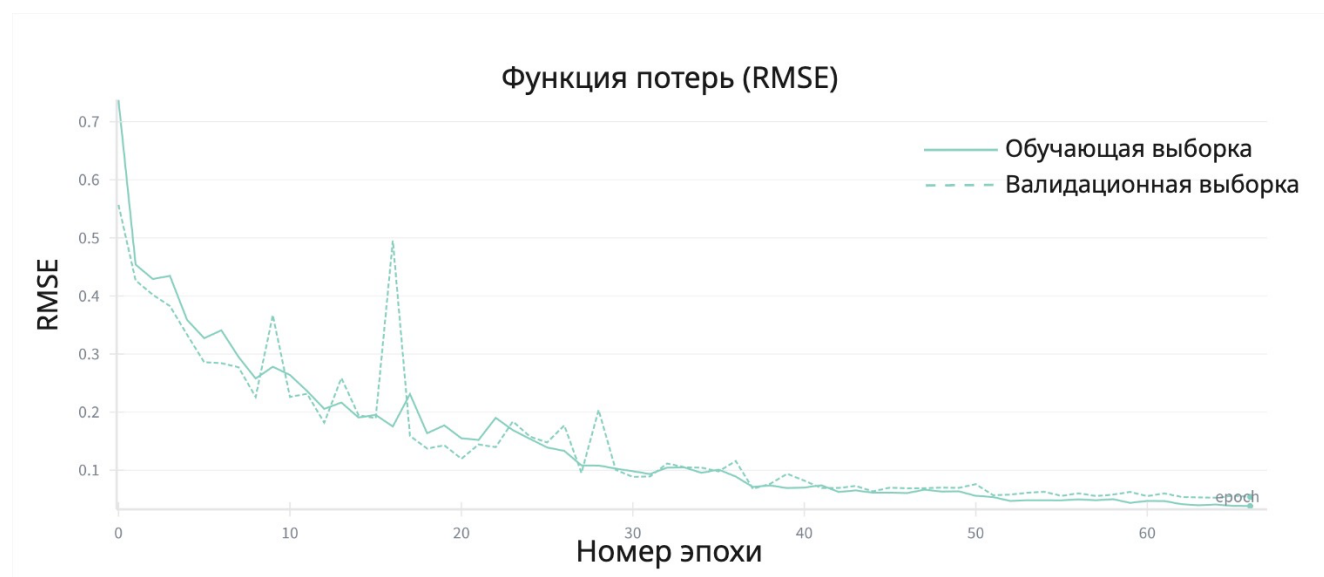


Рисунок 2.10 - Динамика функции потерь (RMSE) для обучающей и валидационной выборок при обучении нейронной сети на данных модели MAID2007.

Таким образом, сравнение с MAID2007 показало, что используемая архитектура нейронной сети и выбранная схема обучения способны адекватно восстанавливать известную зависимость дифференциальных сечений от кинематических переменных. Полученный результат важен по двум причинам. Во-первых, он подтверждает, что нейросетевой алгоритм не вносит заметных нефизических искажений даже при обучении на разреженном наборе точек в многомерном пространстве. Во-вторых, он показывает, что сама постановка задачи

регрессии является корректной и позволяет восстанавливать гладкую многомерную структуру сечений, а не только приближать отдельные численные значения.

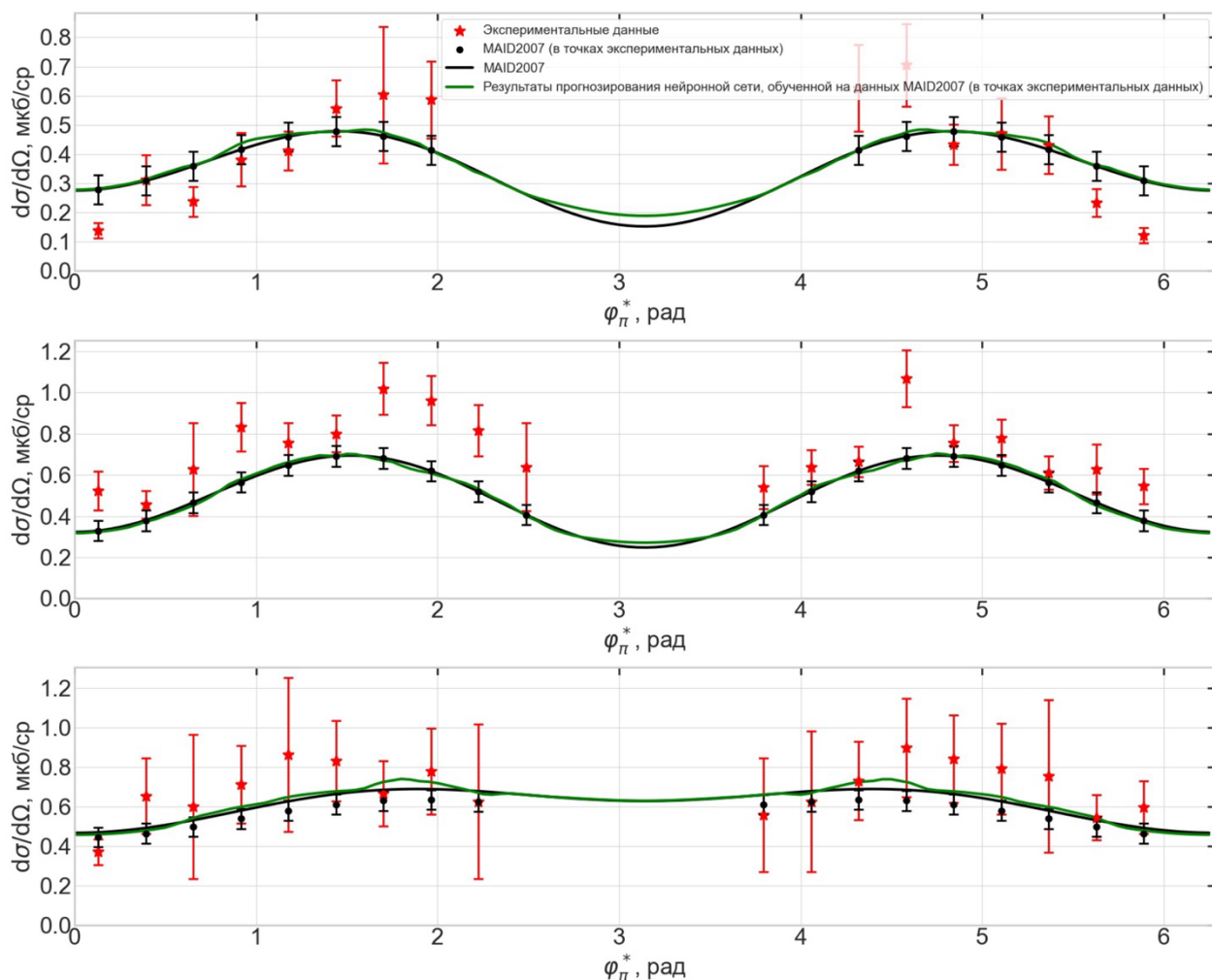


Рисунок 2.11 - Сравнение зависимости дифференциального сечения $\frac{d\sigma}{d\Omega}$ от угла ϕ_π^* для экспериментальных данных, модели MAID2007 и нейронной сети, обученной на прогнозах модели MAID2007. Экспериментальные значения показаны красными точками со статистическими погрешностями, значения прогнозирования модели MAID2007 - черными кривыми, а результаты нейронной сети - зелеными кривыми. Графики соответствуют следующим кинематическим условиям (сверху вниз):

$$E = 5.75 \text{ ГэВ}, W = 1.23 \text{ ГэВ}, Q^2 = 2.915 \text{ ГэВ}^2, \cos\theta_\pi^* = -0.7,$$

$$E = 5.75 \text{ ГэВ}, W = 1.23 \text{ ГэВ}, Q^2 = 2.915 \text{ ГэВ}^2, \cos\theta_\pi^* = -0.1,$$

$$E = 5.75 \text{ ГэВ}, W = 1.745 \text{ ГэВ}, Q^2 = 2.05 \text{ ГэВ}^2, \cos\theta_\pi^* = 0.3.$$

Поэтому сравнение с MAID2007 можно рассматривать как независимую валидацию предложенного подхода и как дополнительный аргумент в пользу применимости нейронной сети для интерполяции наблюдаемых реакции электророждения одиночного пиона в резонансной области.

2.4.3 Оценка статистической устойчивости модели методом реплик

Валидация модели методом реплик используется для оценки ее устойчивости к статистическим флуктуациям экспериментальных данных, а также для определения интервалов неопределенности прогнозируемых значений дифференциальных сечений. Метод реплик представляет собой широко применяемый подход к анализу неопределенностей и устойчивости моделей машинного обучения. Его ключевая идея заключается в формировании ансамбля статистически эквивалентных псевдовыборок данных путем флуктуации исходных измерений в пределах их экспериментальных погрешностей. На каждой из таких выборок обучается независимая модель, после чего разброс полученных прогнозов используется для количественной оценки неопределенностей.

Такой подход позволяет избежать жестких предположений о распространении ошибок и успешно применяется в физике высоких энергий [34], в частности в рамках подхода NNPDF для определения партонных распределений [35, 36].

Процедура валидации методом реплик включает несколько основных этапов:

1. обучение генераторной нейросетевой модели, воспроизводящей экспериментальные значения дифференциальных сечений;
2. построение Монте-Карло реплик целевых значений;
3. обучение ансамбля нейронных сетей на сформированных репликах;

4. анализ отклонений прогнозируемых значений как функций кинематических переменных.

2.4.3.1 Обучение генераторной модели реплик

Здесь, как и в предыдущих главах, рассматриваются экспериментальные данные реакции (2.1). Набор данных неизменно содержит 98 027 измеренных значений дифференциальных сечений, которые в рамках данного подхода удобно обозначить как $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Exp}}$. Генераторная нейронная сеть обучается на 90% экспериментального набора данных с использованием той же архитектуры и стратегии оптимизации, которые применялись в предыдущих разделах работы. Оставшиеся 10% экспериментальных данных используются для контроля качества модели: 5% на этапе валидации и 5% для финальной оценки MAE. Кривые обучения и валидации для генераторной модели приведены на рис. 2.12.

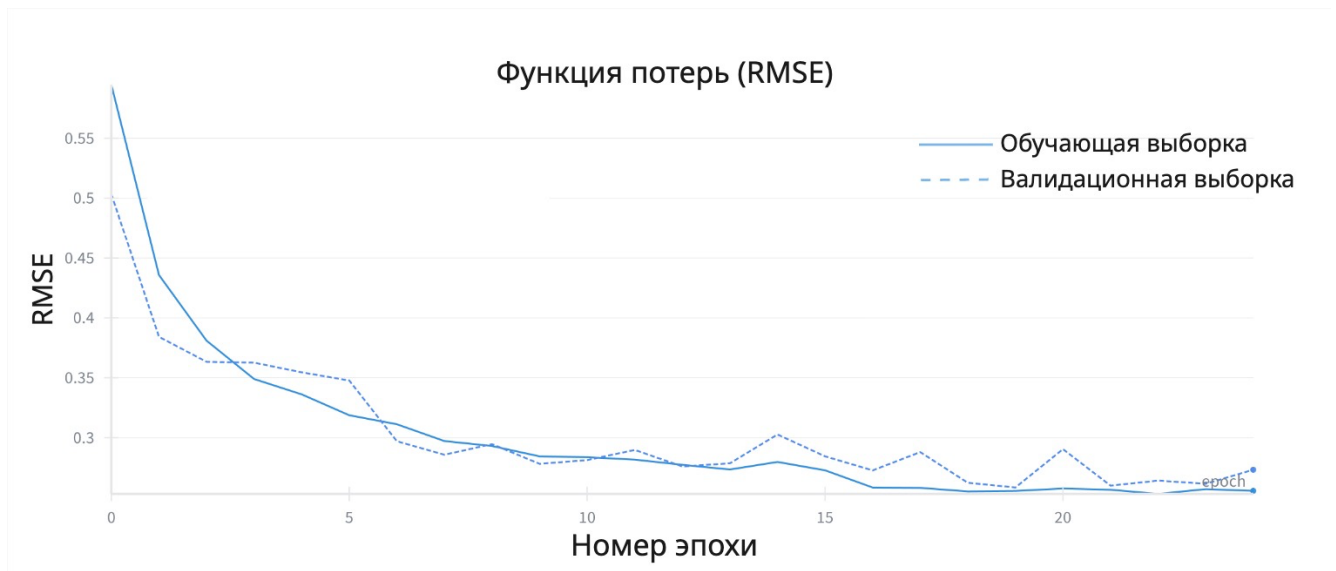


Рисунок 2.12 - Динамика функции потерь (RMSE) для обучающей и валидационной выборок при обучении генераторной нейронной сети в методе реплик.

После обучения сеть формирует для каждой экспериментальной точки прогнозируемое значение $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Gen}}$, которое используется в качестве центрального

значения при генерации реплик. Для ускорения процедуры обучения параметр ранней остановки был уменьшен с 10 до 3 эпох. Это привело лишь к умеренному снижению точности модели, при котором значение ошибки на валидационной выборке составило $MAE \approx 0.09$ мкб/ср.

2.4.3.2 Построение реплик

После обучения генераторной модели для каждой экспериментальной точки формируется ансамбль статистических реплик значений дифференциальных сечений. Каждая экспериментальная точка характеризуется набором кинематических переменных $(E, W, Q^2, \cos \theta_\pi^*, \phi_\pi^*)$, измеренным значением дифференциального сечения $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Exp}}$ и соответствующей экспериментальной неопределенностью $\varepsilon\left(\frac{d\sigma}{d\Omega}\right)_{\text{Exp}}$. Для каждой точки строится нормальное распределение со средним значением $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Gen}}$ и стандартным отклонением $\varepsilon\left(\frac{d\sigma}{d\Omega}\right)_{\text{Exp}}$. Из этого распределения независимо генерируется набор из 100 реплик $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Rep}}^r$, где $r = 1, \dots, 100$. Данная процедура выполняется независимо для всех экспериментальных точек набора данных.

2.4.3.3 Обучение ансамбля реплик-моделей

На следующем этапе обучается ансамбль из 100 независимых нейронных сетей, каждая из которых использует в качестве целевой переменной одну из построенных реплик $\left(\frac{d\sigma}{d\Omega}\right)_{\text{Rep}}^r$. Архитектура сети и все гиперпараметры обучения остаются неизменными. Кривые обучения и валидации для всех реплик-моделей представлены на рис. 2.13.

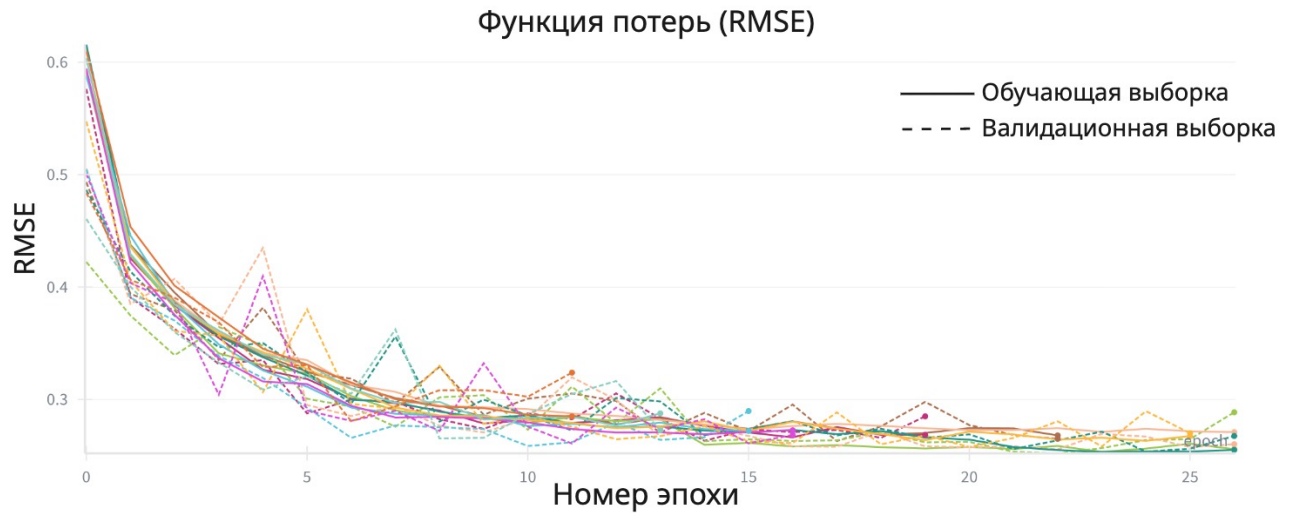


Рисунок 2.13 - Кривые функций потерь (RMSE) для обучающей (сплошные линии) и валидационной (пунктирные линии) выборки, полученные при обучении ансамбля из 100 реплик-моделей. Данный рисунок предназначен не для анализа отдельных траекторий обучения, а для иллюстрации общего характера сходимости и устойчивости алгоритма в целом: несмотря на значительные статистические флуктуации на ранних эпохах, все модели демонстрируют согласованное уменьшение функции потерь и выход на сопоставимый уровень RMSE на стадии сходимости.

Для всех реплик-моделей значение ошибки на валидационной выборке оказалось близким к $MAE \approx 0.09$ мкб/ср, что согласуется с точностью генераторной модели.

2.4.3.4 Анализ результатов метода реплик

После обучения ансамбля нейронных сетей для каждой точки фазового пространства можно вычислить среднее значение прогнозируемого дифференциального сечения и соответствующий интервал неопределенности

$$\Delta_{Rep}^r = \frac{\left(\frac{d\sigma}{d\Omega}\right)_{Rep}^r - \left(\frac{d\sigma}{d\Omega}\right)_{Gen}}{\left(\frac{d\sigma}{d\Omega}\right)_{Gen}}. \quad (2.15)$$

Пример зависимости дифференциального сечения $\frac{d\sigma}{d\Omega}$ от азимутального угла ϕ_π^* в фиксированной области фазового пространства показан на рис. 2.14. Хорошее согласие экспериментальных точек со средней кривой и их расположение внутри полосы неопределенности свидетельствуют о корректном воспроизведении угловой зависимости дифференциального сечения и устойчивости модели.

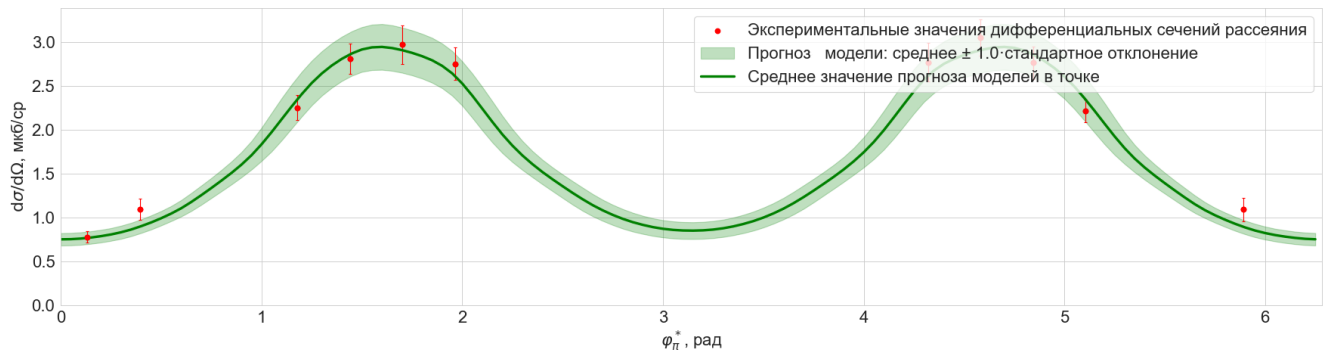


Рисунок 2.14 - Зависимость дифференциального сечения рассеяния $\frac{d\sigma}{d\Omega}$ от азимутального угла ϕ_π^* в фиксированной области фазового пространства. Красными точками с вертикальными погрешностями показаны экспериментальные значения дифференциальных сечений. Сплошная зеленая линия соответствует среднему значению прогнозирования, полученному усреднением по ансамблю независимых нейронных сетей, обученных методом реплик. Заштрихованная область вокруг средней кривой отражает интервал неопределенности, равный одному стандартному отклонению ($\pm 1\sigma$), и характеризует чувствительность прогнозирования модели к статистическим флуктуациям входных данных. Хорошее согласие экспериментальных точек со средней кривой и их расположение внутри полосы неопределенности свидетельствуют о стабильности модели и корректном воспроизведении угловой зависимости дифференциального сечения.

Дополнительно был проведен анализ распределений относительных отклонений, полученных методом реплик, в зависимости от основных кинематических переменных реакции: $\cos \theta_\pi^*$, ϕ_π^* , Q^2 , W . Полученные распределения демонстрируют отсутствие систематических смещений и подтверждают устойчивость предложенной нейросетевой модели в различных областях кинематического фазового пространства (рис. 2.15).

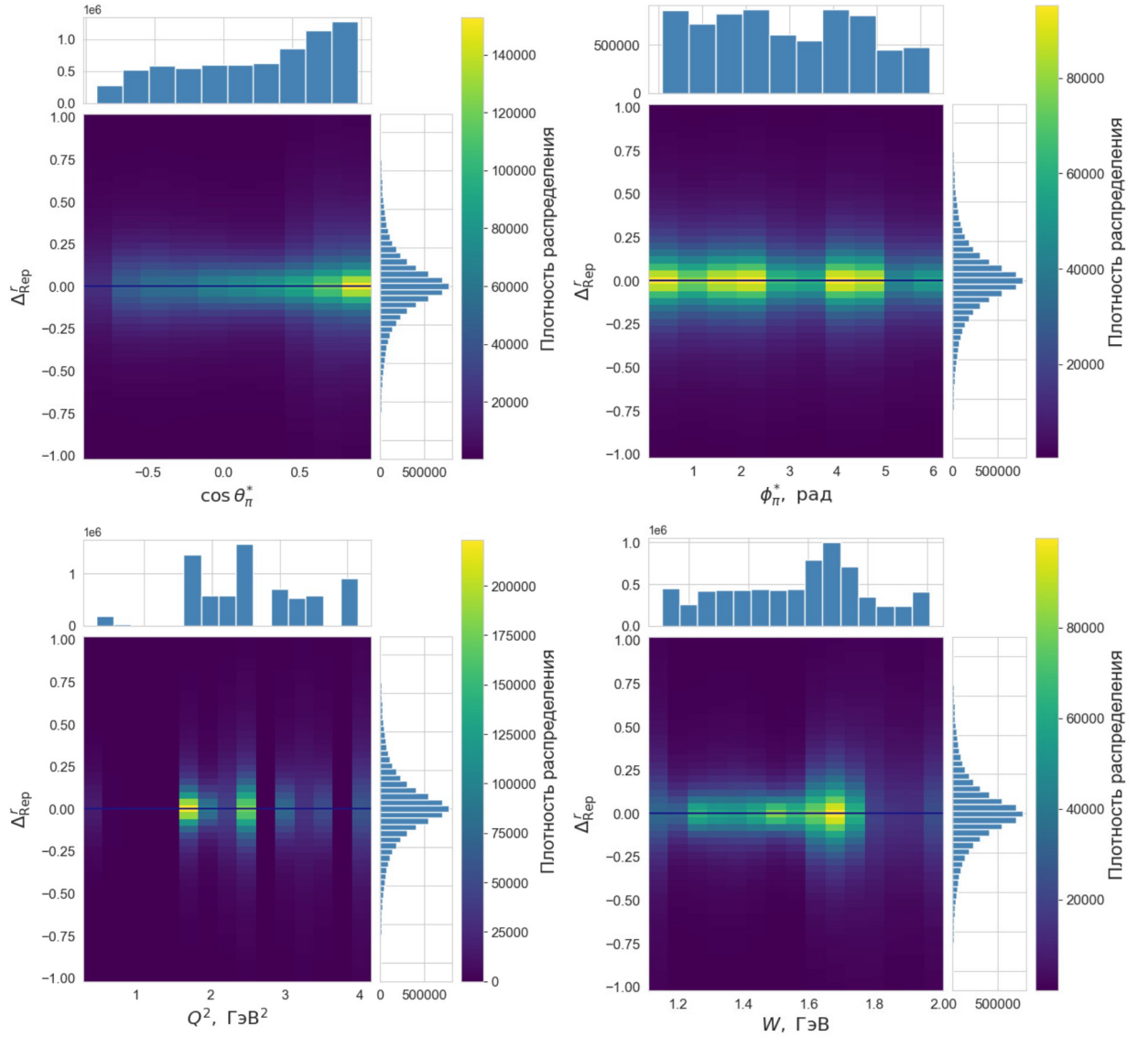


Рисунок 2.15 - Распределения относительного отклонения Δ_{Rep}^r , полученного методом реплик, в зависимости от основных кинематических переменных реакции: косинуса полярного угла эмиссии пиона $\cos \theta_{\pi}^*$, азимутального угла ϕ_{π}^* , квадрата переданного четырехимпульса Q^2 и инвариантной массы адронной системы W . Цветовая шкала отражает плотность распределения событий в каждой ячейке двумерного пространства $(x, \Delta_{\text{Rep}}^r)$, тогда как верхние и боковые гистограммы показывают соответствующие одномерные маргинальные распределения по каждой из переменных. Горизонтальная линия $\Delta_{\text{Rep}}^r = 0$ соответствует отсутствию систематического смещения между прогнозированием ансамбля моделей и центральным значением. Концентрация распределений вблизи нуля для всех рассмотренных кинематических переменных свидетельствует о стабильности и однородности прогнозирования нейросетевой модели во всем диапазоне исследуемых параметров.

3 Оценка сечений эксклюзивных каналов электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$

В данной главе описывается метод оценки сечений электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах в области инвариантных масс конечных адронов $M_K + M_Y < W < 2.65$ ГэВ (M_K и M_Y - массы каона и гиперона) и квадратов четырехимпульса виртуальных фотонов, т. н. виртуальностей фотона $0 < Q^2 < 5 \text{ ГэВ}^2$ на основе экспериментальных данных по сечениям этих эксклюзивных каналов, измеренных на детекторе CLAS в национальной исследовательской лаборатории имени Т. Джефферсона. Описывается процедура интерполяции и экстраполяции структурных функций каналов электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ по углу эмиссии каона в системе центра масс конечных адронов с применением различных алгоритмов прогнозирования поведения исследуемых наблюдаемых. В данной главе также описана реализация алгоритмов для оценки дифференциальных сечений данных каналов с учетом статистических и систематических неопределенностей. В процессе исследования был разработан комплекс программ для оценки дифференциальных сечений и структурных функций на основе языков программирования C++ и Python. Для интерактивной работы с программным комплексом был создан веб-сайт, позволяющий анализировать одномерные и двумерные зависимости структурных функций и дифференциальные сечения. Оценка дифференциальных сечений каналов электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ необходима для извлечения структурной функции $\sigma_{LT'}$ из данных по поляризационной асимметрии реакций электророждения этих конечных состояний продольно поляризованными электронами. Полученные результаты также важны для развития реалистических Монте-Карло генераторов событий при планировании будущих экспериментов и для оценки эффективности регистрации конечных частиц при извлечении сечений реакций из экспериментальных данных [37].

3.1 Экспериментальные данные по электровозбуждению нуклонов

Изучение эксклюзивных каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах представляет собой важную часть исследований спектра и структуры возбужденных состояний нуклона (N^*) или нуклонных резонансов в области масс > 1.6 ГэВ [38]. Измерения на детекторе CLAS обеспечили значительную часть имеющейся в мире информации о большинстве эксклюзивных каналов электророждения мезонов в области энергий возбуждения нуклонных резонансов $W < 2.0$ ГэВ, $Q^2 < 5.0$ ГэВ² [39].

Конечная адронная система	W , ГэВ	Q^2 , ГэВ ²	Измеренные наблюдаемые
π^+n	1.1 - 1.38	0.16 - 0.36	$d\sigma/d\Omega$
	1.1 - 1.55	0.3 - 0.6	$d\sigma/d\Omega$
	1.1 - 1.7	1.7 - 4.5	$d\sigma/d\Omega, A_b$
	1.6 - 2.0	1.8 - 4.5	$d\sigma/d\Omega$
π^0p	1.1 - 1.38	0.16 - 0.36	$d\sigma/d\Omega$
	1.1 - 1.68	0.4 - 1.8	$d\sigma/d\Omega, A_b, A_t, A_{bt}$
	1.1 - 1.39	3.0 - 6.0	$d\sigma/d\Omega$
ηp	1.5 - 2.3	0.2 - 3.1	$d\sigma/d\Omega$
$K^+\Lambda$	$M(K^+\Lambda) - 2.6$	1.4 - 3.9	$d\sigma/d\Omega$
		0.7 - 5.4	P^0, P'
$K^+\Sigma^0$	$M(K^+\Sigma^0) - 2.6$	1.4 - 3.9	$d\sigma/d\Omega$
		0.7 - 5.4	P'
$\pi^+\pi^-p$	1.3 - 1.6	0.2 - 0.6	9 одномерных дифференциальных сечений
	1.4 - 2.1	0.5 - 1.5	
	1.4 - 2.0	2.0 - 5.0	
A_b — асимметрия пучка, A_t — асимметрия мишени, A_{bt} — двойная асимметрия пучок-мишень; P^0, P' — поляризации гиперона отдачи			

Таблица 1 - Доступные экспериментальные данные электророждения мезонов на протоне.

Кинематические области и соответствующие наблюдаемые, полученные в реакциях электророждения мезонов на протонах, представлены в Таблице 1. Измеренные наблюдаемые получены с почти полным охватом ($\approx 4\pi$) угла вылета

конечного каона в системе центра масс конечных адронов, что критически важно для надежного извлечения амплитуд электровозбуждения нуклонных резонансов. Результаты измерений сечений и поляризационных наблюдаемых реакций электророждения мезонов находятся в CLAS Physics Database.

Каналы	Возбужденные нуклонные состояния	Q^2 , ГэВ ²
$\pi^0 p \pi^+ n$	$\Delta(1232)3/2^+$	0.16 – 6.0
	$N(1440)1/2^+ \quad N(1520)3/2^- \quad N(1535)1/2^-$	0.3 – 4.16
$\pi^+ n$	$N(1675)5/2^- \quad N(1680)5/2^+ \quad N(1710)1/2^+$	1.6 – 4.5
ηp	$N(1535)1/2^-$	0.2 – 2.9
$\pi^+ \pi^- p$	$N(1440)1/2^+ \quad N(1520)3/2^-$	0.25 – 1.5
	$\Delta(1620)1/2^- \quad N(1650)1/2^- \quad N(1680)5/2^+$	0.5 – 1.5
	$\Delta(1700)3/2^- \quad N(1720)3/2^+ \quad N'(1720)3/2^+$	

Таблица 2 - Доступные амплитуды электровозбуждения резонансов в различных каналах из данных детектора CLAS.

В настоящее время информация об амплитудах электровозбуждения (N^*) получена преимущественно из реакций электророждения $N\pi$ и $\pi^+\pi^-p$ [38, 40]. Сводка имеющихся результатов представлена в Таблице 2. Результаты по амплитудам электровозбуждения N^* , полученные из этих двух каналов с различными нерезонансными вкладами, находятся в хорошем согласии. Данный успех позволил подтвердить надежность моделей реакций, развитых в лаборатории им. Т. Джефферсона для извлечения амплитуд электровозбуждения N^* из реакций электророждения $N\pi$ и $\pi^+\pi^-p$ на протонах.

Значительная часть возбужденных состояний нуклона с массами выше 1.6 ГэВ распадается преимущественно с испусканием двух пионов. Экспериментальные результаты по амплитудам электровозбуждения этих резонансов были получены при исследовании двухпионного канала [41]. Сравнение этих результатов с данными, полученными из других эксклюзивных каналов, позволяет удостовериться в надежности извлечения амплитуд электровозбуждения N^* . Однако однопионный канал не обладает достаточной чувствительностью к амплитудам электровозбуждения N^* , основными каналами распада которых

являются двухпионные состояния. Независимое извлечение амплитуд электровозбуждения этих резонансов из реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ позволит независимо проверить достоверность извлечения амплитуд электровозбуждения резонансов N^* .

Важной частью экспериментов по извлечению сечений реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ является оценка эффективности регистрации конечных продуктов реакции. Для получения этих оценок очень часто прибегают к использованию Монте-Карло генераторов событий, основанных на экспериментальных данных по сечениям реакций $K^+ \Lambda$ и $K^+ \Sigma^0$ в широкой области кинематических переменных W , Q^2 и угла эмиссии конечного каона в системе центра масс конечных адронов. Оценка эффективности регистрации конечных продуктов реакции также необходима для извлечения структурной функции $\sigma_{LT'}$ из измерений асимметрии пучка в реакциях с продольно-поляризованными электронами.

В работе развит метод оценки дифференциальных сечений реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах в области $W \in [M_K + M_Y, 2.65]$ ГэВ и $Q^2 \in [0, 5]$ ГэВ² из данных, полученных на детекторе CLAS. Сравнение с имеющимися результатами подтвердило надёжность метода; также были оценены неопределённости прогнозируемых значений сечений.

3.2 Методы оценки дифференциальных сечений электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах из экспериментальных данных CLAS

Оценка дифференциальных сечений реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах выполнена на основе данных детектора CLAS по структурным

функциям $\frac{d\sigma_U}{d\Omega_K}(E, W, Q^2, \cos \theta)$, $\frac{d\sigma_{TT}}{d\Omega_K}(W, Q^2, \cos \theta)$ и $\frac{d\sigma_{LT}}{d\Omega_K}(W, Q^2, \cos \theta)$,
извлеченных из измеренных дифференциальных сечений этих реакций [42, 43].

3.2.1 Формализм описания реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах

В приближении однофотонного обмена указанные выше структурные функции связаны с дифференциальными сечениями электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах под действием виртуальных фотонов $\frac{d\sigma_{\gamma^*}}{d\Omega_K}$ следующим образом [18] (аналогично (2.7) с точностью до замены ϕ_π^* на ϕ_K^*):

$$\frac{d\sigma_{\gamma^*}}{d\Omega_K} = \frac{d\sigma_T}{d\Omega_K} + \varepsilon \frac{d\sigma_L}{d\Omega_K} + \varepsilon \frac{d\sigma_{TT}}{d\Omega_K} \cos 2\phi_K^* + \sqrt{\varepsilon(1+\varepsilon)} \frac{d\sigma_{LT}}{d\Omega_K} \cos \phi_K^*, \quad (3.1)$$

где ϕ_K^* — азимутальный угол эмиссии каона в системе центра масс $\gamma^* p$ относительно плоскости, в которой лежат трехимпульсы падающего и рассеянного электронов в лабораторной системе. Первые два члена в правой части (3.1) описывают динамику электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ на протонах под действием виртуальных фотонов в состояниях с поперечной $\frac{d\sigma_T}{d\Omega_K}$ и продольной $\frac{d\sigma_L}{d\Omega_K}$ поляризациями. Их сумма представляет собой так называемую неполяризованную структурную функцию $\frac{d\sigma_U}{d\Omega_K}$:

$$\frac{d\sigma_U}{d\Omega_K} = \frac{d\sigma_T}{d\Omega_K} + \varepsilon \frac{d\sigma_L}{d\Omega_K}, \quad (3.2)$$

где ε — поперечная поляризация виртуального фотона, определяемая кинематикой рассеяния электронов на протонах:

$$\varepsilon = \left(1 + 2 \frac{|\vec{q}|^2}{Q^2} t g^2 \frac{\theta_{e'}}{2}\right)^{-1}, \quad (3.3)$$

где $|\vec{q}|$ — модуль трехимпульса виртуального фотона, $\theta_{e'}$ — полярный угол рассеяния электрона в лабораторной системе координат. Поскольку виртуальные

фотоны, формируемые в процессе рассеяния электронов на протонах, оказываются поляризованными, то появляется зависимость дифференциальных сечений от угла ϕ_K^* , описываемая выражением (3.1). Структурная функция $\frac{d\sigma_{TT}}{d\Omega_K}(W, Q^2, \cos \theta_K^*)$ описывает электророждение $K^+ \Lambda$ и $K^+ \Sigma^0$ виртуальными фотонами в состояниях со смешанными поперечными поляризациями. Структурная функция $\frac{d\sigma_{LT}}{d\Omega_K}(W, Q^2, \cos \theta_K^*)$ описывает данную реакцию под действием фотонов в состояниях со смешанными продольной и поперечной поляризациями.

Измеренные дифференциальные сечения реакций электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ связаны с сечениями реакций под действием виртуальных фотонов $\frac{d\sigma_{\gamma^*}}{d\Omega_K}$ соотношением:

$$\frac{d\sigma}{dQ^2 dW d\Omega_K} = \Gamma_{\gamma^*} \frac{d\sigma_{\gamma^*}}{d\Omega_K}. \quad (3.4)$$

Поток виртуальных фотонов Γ_{γ^*} определяется кинематикой рассеяния электронов. Поток виртуальных фотонов не соответствует реальному потоку частиц; он определяется в рамках используемого формализма и зависит от выбранного соглашения. Использовалось соглашение о потоке виртуальных фотонов (так называемое *hand convention*), для которого были получены сечения электророждения мезонов на протонах из измерений на установке CLAS:

$$\Gamma_{\gamma^*} = \frac{\alpha}{4\pi} \frac{W}{M_p^2 E^2} \frac{W^2 - M_p^2}{Q^2} \frac{1}{1 - \epsilon}, \quad (3.5)$$

где M_p — масса протона, $\alpha = \frac{1}{137}$.

В работе используются экспериментальные данные CLAS с неполяризованной протонной мишенью при энергиях пучка электронов 2.567, 4.056 и 5.499 ГэВ [42, 43]. Кинематические области по W и Q^2 , где получены сечения электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$, показаны на рис. 3.1.

Из экспериментальных данных детектора CLAS были получены неполяризованные структурные функции $\frac{d\sigma_U}{d\Omega_K}$, описываемые суммой (3.2).

Неполяризованные структурные функции зависят от энергии пучка электронов, определяющей параметр поляризации виртуального фотона ε (3.3). Оценка отношения продольной $\frac{d\sigma_L}{d\Omega_K}$ и поперечной $\frac{d\sigma_T}{d\Omega_K}$ структурных функций выполнена в приближении [44]:

$$\frac{d\sigma_L}{d\Omega_K} : \frac{d\sigma_T}{d\Omega_K} = 0.2 . \quad (3.6)$$

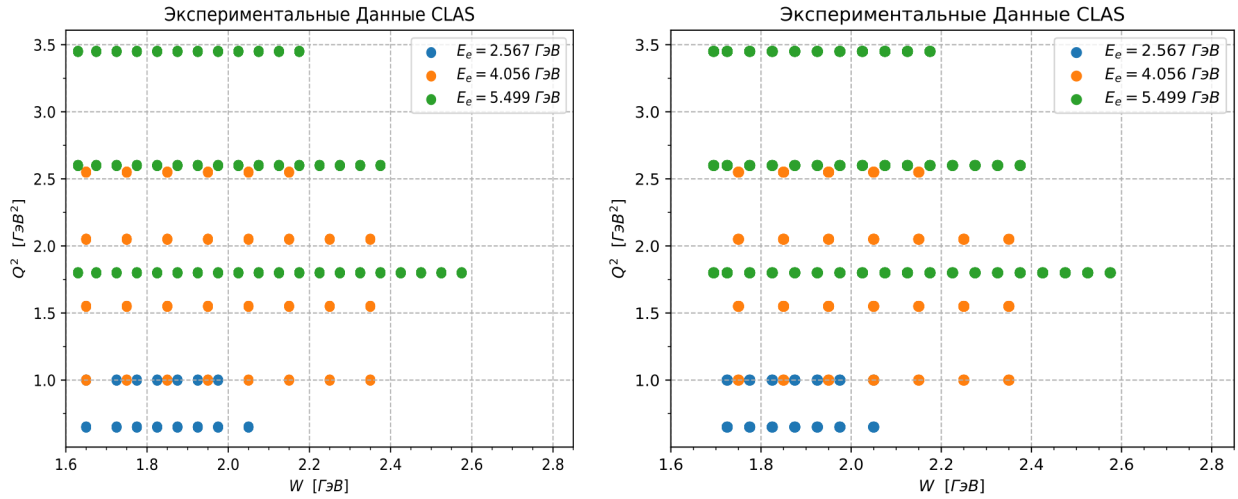


Рисунок 3.1 - Кинематические области по инвариантной массе конечной системы адронов W и виртуальностей фотона Q^2 , в которых получены данные по сечениям каналов $K^+\Lambda$ (слева) и $K^+\Sigma^0$ (справа) из измерений на детекторе CLAS [45, 46]. Синими точками показаны значения, полученные при энергии пучка $E_e = 2.567$ ГэВ, оранжевыми - при $E_e = 4.056$ ГэВ, зелеными - при $E_e = 5.499$ ГэВ.

Используя (3.2) и (3.6), получаем для структурных функций $\frac{d\sigma_T}{d\Omega_K}(W, Q^2, \cos \theta_K^*)$ и $\frac{d\sigma_L}{d\Omega_K}(W, Q^2, \cos \theta_K^*)$:

$$\frac{d\sigma_T}{d\Omega_K} = \frac{\frac{d\sigma_U}{d\Omega_K}}{1+0.2\varepsilon}, \quad \frac{d\sigma_L}{d\Omega_K} = \frac{\frac{d\sigma_U}{d\Omega_K}}{5+\varepsilon} . \quad (3.7)$$

Структурные функции $\frac{d\sigma_T}{d\Omega_K}$ и $\frac{d\sigma_L}{d\Omega_K}$ не зависят от энергии пучка электронов.

Таким образом, результаты по структурным функциям $\frac{d\sigma_T}{d\Omega_K}$ и $\frac{d\sigma_L}{d\Omega_K}$ позволяют восстановить дифференциальные сечения реакций электроорождения для любой энергии пучка электронов в пределах исследуемой кинематической области.

3.3 Интерполяция эксклюзивных структурных функций

В режиме интерполяции были вычислены оценки дифференциальных сечений электророждения $K^+ \Lambda$ и $K^+ \Sigma^0$ в кинематической области $W \in [M_K + M_Y, 2.65]$ ГэВ, $Q^2 \in [0, 5]$ ГэВ², $\cos \theta_K^* \in [-1, 1]$ из данных CLAS без использования специфических, модельных, априорных предположений о динамике этих реакций. Более общей задачей является задача восстановления значений структурных функций во всем указанном сегменте фазового пространства. Так как исходные данные были получены на дискретной сетке величин $(W, Q^2, \cos \theta_K^*)$, необходимо провести интерполяцию эксклюзивных структурных функций на сетку кинематических переменных, где будут оцениваться дифференциальные сечения.

Зависимости структурных функций от полярного угла эмиссии каона в системе центра масс конечных адронов описывались разложением по полиномам Лежандра:

$$\begin{aligned} \frac{d\sigma_{T/L}}{d\Omega_K} &= A_{T/L} \cdot P_0(\cos \theta_K^*) + B_{T/L} \cdot P_1(\cos \theta_K^*) + C_{T/L} \cdot P_2(\cos \theta_K^*) + D_{T/L} \cdot \\ &\quad P_3(\cos \theta_K^*) + E_{T/L} \cdot P_4(\cos \theta_K^*), \\ \frac{d\sigma_{LT}}{d\Omega_K} &= \sin \theta_K^* [A_{LT} \cdot P_0(\cos \theta_K^*) + B_{LT} \cdot P_1(\cos \theta_K^*) + C_{LT} \cdot P_2(\cos \theta_K^*) + D_{LT} \cdot \\ &\quad P_3(\cos \theta_K^*) + E_{LT} \cdot P_4(\cos \theta_K^*)], \\ \frac{d\sigma_{TT}}{d\Omega_K} &= \sin^2 \theta_K^* [A_{TT} \cdot P_0(\cos \theta_K^*) + B_{TT} \cdot P_1(\cos \theta_K^*) + C_{TT} \cdot P_2(\cos \theta_K^*) + D_{TT} \cdot \\ &\quad P_3(\cos \theta_K^*) + E_{TT} \cdot P_4(\cos \theta_K^*)]. \end{aligned} \quad (3.8)$$

Коэффициенты разложения по полиномам Лежандра (3.8), извлеченные из экспериментальных данных, были интерполированы по переменным W и Q^2 с использованием линейной аппроксимации.

Сравнение с экспериментальными данными (рис. 3.2) показало, что разложение по пяти полиномам Лежандра обеспечивает хорошее описание во всем кинематическом диапазоне интерполяции. Неопределенности структурных

функций получены из неопределенностей экспериментальных данных путем интерполяции кубическим сплайном. Данная процедура позволяет получить структурные функции на сетке (W, Q^2) в области фазового пространства, где имеются экспериментальные данные (рис. 3.2).

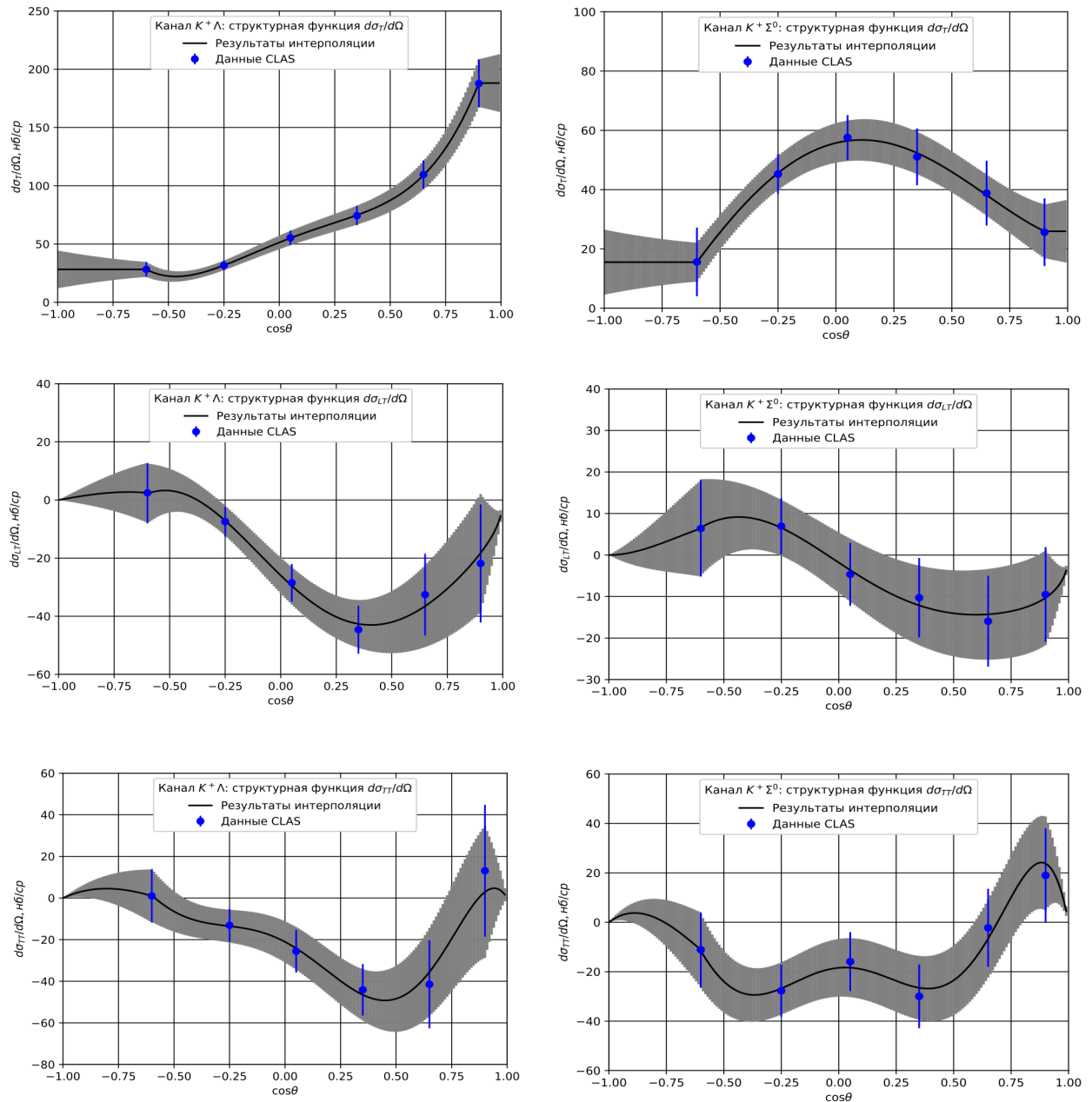


Рисунок 3.2 - Сравнение результатов интерполяции структурных функций (черная линия) и экспериментальных данных для канала

$K^+ \Lambda$ (слева) и $K^+ \Sigma^0$ (справа) при $W = 1.95$ ГэВ, $Q^2 = 1.0$ ГэВ². Синим показаны экспериментальные данные CLAS [45, 46] и их неопределенности. Неопределенности интерполяции представлены серой областью.

Описанные процедуры позволяют оценить дифференциальные сечения на основе структурных функций $\frac{d\sigma_T}{d\Omega_K}$, $\frac{d\sigma_L}{d\Omega_K}$, $\frac{d\sigma_{LT}}{d\Omega_K}$ и $\frac{d\sigma_{TT}}{d\Omega_K}$ на исследуемой сетке кинематических переменных W , Q^2 , $\cos(\theta_K^*)$ там, где возможно использовать интерполяцию структурных функций. В данной работе также дается описание процедуры экстраполяции структурных функций по переменным W , Q^2 , $\cos(\theta_K^*)$, которая позволяет восстановить их значения в области $W < 2.65$ ГэВ. Расширение диапазона оценки структурных функций по W до 2.65 ГэВ необходимо для экспериментов на детекторе CLAS12 по поиску так называемых гибридных барионов с одетыми глюонами в качестве активной компоненты их структуры наряду с тремя одетыми кварками [45].

3.4 Экстраполяция структурных функций в каналах электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах

Данные CLAS по структурным функциям электророждения $K^+\Lambda$ и $K^+\Sigma^0$ представлены преимущественно в кинематической области $W < 2.4$ ГэВ и $Q^2 < 3.6$ ГэВ² (рис. 3.1). В настоящей работе выполнена экстраполяция структурных функций в области $W < 2.65$ ГэВ и $Q^2 < 5$ ГэВ². Данный процесс экстраполяции позволяет создать модели Монте-Карло генераторов событий для перечисленных каналов электророждения, используя прогнозируемые значения сечения в области $W < 2.65$ ГэВ и $Q^2 < 5$ ГэВ²; эти значения затем могут применяться для надежного определения эффективности регистрации $K^+\Lambda$ и $K^+\Sigma^0$ событий в экспериментах на детекторе CLAS12. Сечения этих каналов играют важную роль в исследованиях спектра и структуры нуклонных резонансов, в особенности в поиске гибридных барионов и "missing"-резонансов [45, 46].

3.4.1 Экстраполяция структурных функций по W и Q^2

Полученные из данных детектора CLAS структурные функции электророждения $K^+\Lambda$ и $K^+\Sigma^0$ охватывают преимущественно область $W < 2.4$ ГэВ. В то же время, при $Q^2 = 1.8$ ГэВ² они получены в области W до 2.57 ГэВ. Экстраполяция структурных функций в область W до 2.575 ГэВ для интервала $Q^2 < 3.6$ ГэВ² выполнена в предположении, что при любом значении Q^2 отношение структурных функций в соседних значениях по W : W_j и W_{j+1} совпадает с отношением, определенным из измерений при $Q^2 = 1.8$ ГэВ². Это дает возможность экстраполяции структурных функций по W в области W до 2.575 ГэВ и $Q^2 < 3.6$ ГэВ² следующим образом:

$$\begin{aligned} \frac{d\sigma_i}{d\Omega_K}(W_{j+1}, Q_{curr}^2 \text{ ГэВ}^2, \cos(\theta_K^*)) &= \frac{d\sigma_i}{d\Omega_K}(W_{j+1}, Q^2 = 1.8 \text{ ГэВ}^2, \cos(\theta_K^*)) \times \\ \frac{d\sigma_i}{d\Omega_K}(W_j, Q_{curr}^2 \text{ ГэВ}^2, \cos(\theta_K^*)) &: \frac{d\sigma_i}{d\Omega_K}(W_j, Q^2 = 1.8 \text{ ГэВ}^2, \cos(\theta_K^*)), \end{aligned} \quad (3.9)$$

где Q_{curr}^2 – текущее значение виртуальности, $i = U, T, TT, TL$. На рис. 3.3 представлены результаты экстраполяции зависимостей дифференциальных сечений от инвариантной массы W для каналов $K^+\Lambda$ и $K^+\Sigma^0$ в диапазоне $Q^2 > 1.8$ ГэВ² согласно (3.9) в сравнении с экспериментальными данными при $Q^2 = 1.8$ ГэВ² и $\cos(\theta_K^*) = 0.1$ для канала $K^+\Lambda$, $\cos(\theta_K^*) = 0.3$ для канала $K^+\Sigma^0$. Экстраполяция в область $2.57 < W < 2.65$ ГэВ выполнена с использованием квадратичного сплайна.

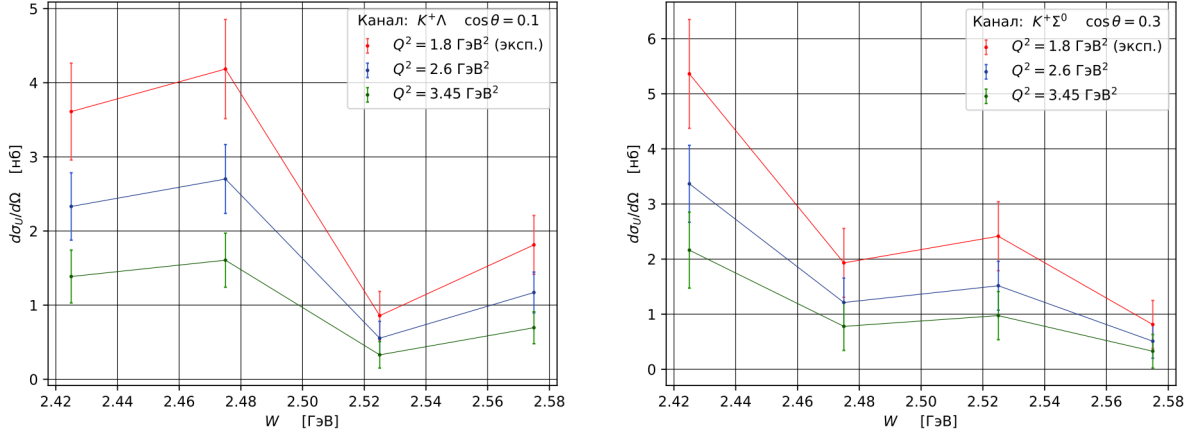


Рисунок 3.3 - Результаты экстраполяции неполяризованных структурных функций для эксклюзивных каналов электроорождения $K^+ \Lambda$ (слева) и $K^+ \Sigma^0$ (справа) на протонах в области $2.40 < W < 2.57$ ГэВ и $1.8 < Q^2 < 3.50$ ГэВ² согласно (3.9) в сравнении с экспериментальными данными [45, 46] при $Q^2 = 1.8$ ГэВ² и $\cos(\theta_K^*) = 0.1$ для канала $K^+ \Lambda$, $\cos(\theta_K^*) = 0.3$ для канала $K^+ \Sigma^0$ при энергии пучка электронов 5.499 ГэВ.

3.4.2 Экстраполяция структурных функций в область

$$3.45 < Q^2 < 5.00 \text{ ГэВ}^2$$

В этой области проводится экстраполяция всех структурных функций из области интерполяции с помощью степенной функции:

$$\frac{d\sigma_i}{d\Omega_K} = A + \frac{B}{Q^2} + \frac{C}{Q^4}. \quad (3.10)$$

Использование зависимости (3.10) основано на предположении, что эволюция эксклюзивных структурных функций описывается теми же степенными зависимостями, что и Q^2 -эволюция моментов инклюзивных структурных функций [47]. Зависимость (3.10) хорошо описывает Q^2 эволюцию структурных функций при $Q^2 > 3.45$ ГэВ², где имеются экспериментальные данные, на которых независимо подбираются параметры в каждом W -интервале. Описание данных и результаты их экстраполяции в область $3.45 < Q^2 < 5$ ГэВ² показаны на рис. 3.4.

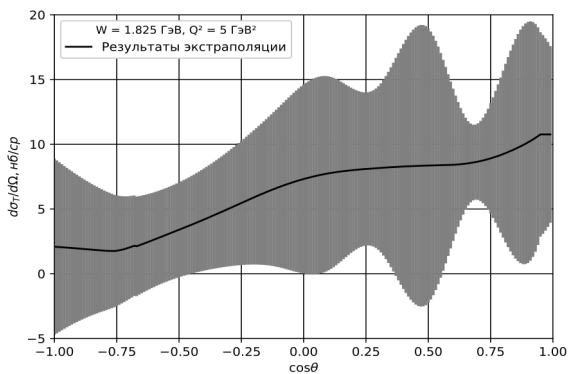
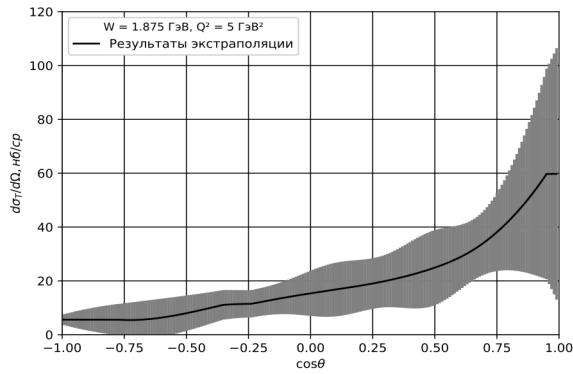
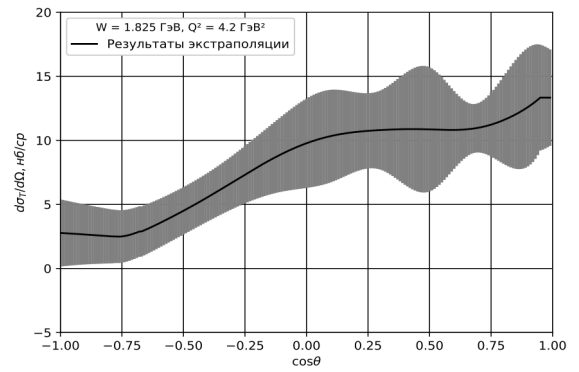
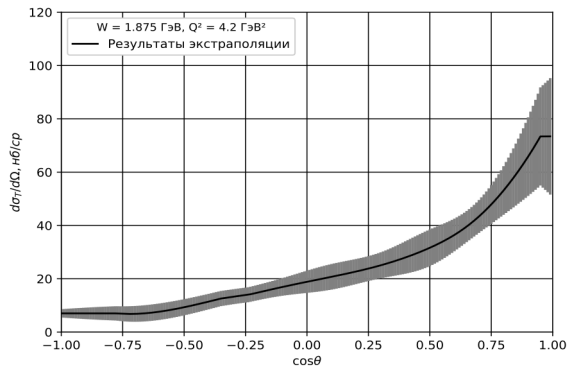
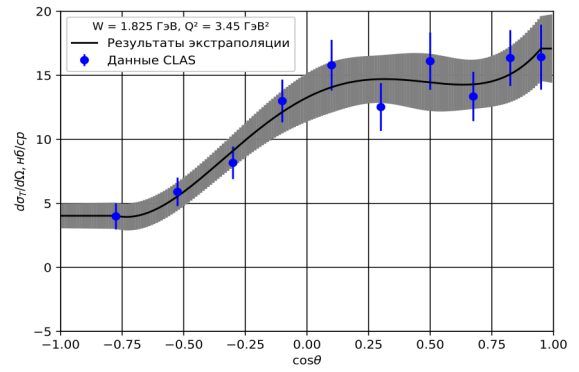
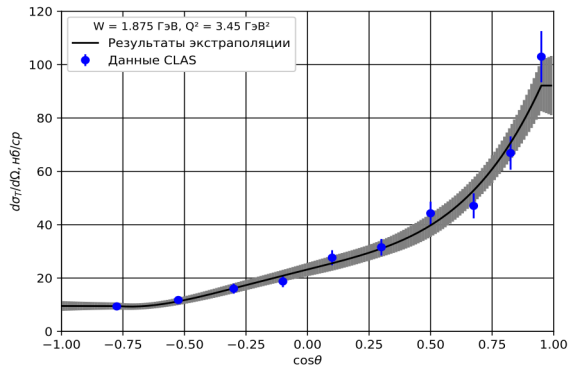


Рисунок 3.4 - Экстраполяция структурной функции $\frac{d\sigma_T}{d\Omega_K}$ в область $Q^2 > 3.45 \text{ ГэВ}^2$: для каналов $K^+\Lambda$ (слева) и $K^+\Sigma^0$ (справа). Экспериментальные данные при $Q^2 = 3.45 \text{ ГэВ}^2$ показаны синими точками и их неопределенностями, а результаты интерполяции показаны черной линией. Результаты экстраполяции показаны черными линиями - при $Q^2 = 4.2 \text{ ГэВ}^2$ и $Q^2 = 5 \text{ ГэВ}^2$. Неопределенности интерполяции и экстраполяции показаны серыми областями.

3.4.3 Экстраполяция структурных функций в области $Q^2 < 0.65 \text{ ГэВ}^2$

В области малых значений Q^2 нет возможности использовать степенную функцию (3.10) для экстраполяции, так как особенность при $Q^2 = 0 \text{ ГэВ}^2$ будет приводить к нефизической расходимости. Разложение (3.10) может быть использовано лишь в области $Q^2 \gg \Lambda_{QCD}^2$. Для экстраполяции структурных функций в области $Q^2 < 0.65 \text{ ГэВ}^2$ использовались ограничения, налагаемые данными по реакциям фоторождения. Измеренные сечения фоторождения $K^+ \Lambda$ и $K^+ \Sigma^0$ в реакциях с неполяризованными фотонами на неполяризованной мишени обеспечивают информацию о структурной функции $\frac{d\sigma_T}{d\Omega_K}$ в фотонной точке при $Q^2 = 0 \text{ ГэВ}^2$. Структурная функция $\frac{d\sigma_T}{d\Omega_K}$ была интерполирована по Q^2 в области $0 < Q^2 < 0.65 \text{ ГэВ}^2$ с использованием полинома:

$$\frac{d\sigma_T}{d\Omega_K} = A \cdot Q^4 + B \cdot Q^2 + C. \quad (3.11)$$

Структурная функция $\frac{d\sigma_{TT}}{d\Omega_K}$ в фотонной точке определялась на основе экспериментальных данных о поляризационной асимметрии пучка фотонов Σ в экспериментах по фоторождению $K^+ \Lambda$ и $K^+ \Sigma^0$ с пучком линейно-поляризованных фотонов на неполяризованной мишени:

$$\frac{d\sigma_{TT}}{d\Omega_K}(W, Q^2 = 0 \text{ ГэВ}^2, \cos(\theta_K^*)) = (-1) \cdot \Sigma \cdot \frac{d\sigma_T}{d\Omega_K}(W, Q^2 = 0 \text{ ГэВ}^2, \cos(\theta_K^*)). \quad (3.12)$$

Следует отметить, что экспериментальные данные по этой наблюдаемой ограничены в интервале по W :

- $W \in [1.72, 2.18] \text{ ГэВ}$ для $K^+ \Lambda$,
- $W \in [1.78, 2.17] \text{ ГэВ}$ для $K^+ \Sigma^0$.

Оценки структурной функции $\frac{d\sigma_{TT}}{d\Omega_K}$ в области $2.17 < W < 2.65 \text{ ГэВ}$ и $Q^2 < 1.8 \text{ ГэВ}^2$ получены в предположении, что при фиксированных значениях Q^2 и $\cos(\theta_K^*)$, отношение $\frac{d\sigma_{TT}}{d\Omega_K} : \frac{d\sigma_T}{d\Omega_K}$ остается постоянным в диапазоне $2.17 < W < 2.65 \text{ ГэВ}$.

Данное предположение соответствует данным CLAS по $K^+\Lambda$ и $K^+\Sigma^0$ фото- и электроррождению. Для каждого значения W определялось отношение $\frac{d\sigma_{TT}}{d\Omega_K} : \frac{d\sigma_T}{d\Omega_K}$, которое усреднялось по W от 2.0 ГэВ до 2.65 ГэВ. После этого структурные функции в области $2.17 < W < 2.65$ ГэВ и $Q^2 < 1.8$ ГэВ² оценивались следующим образом:

$$\frac{d\sigma_{TT}}{d\Omega_K}(W, Q^2, \cos(\theta_K^*)) = \left[\frac{d\sigma_{TT}}{d\Omega_K} : \frac{d\sigma_T}{d\Omega_K} \right](Q^2, \cos(\theta_K^*)) \cdot \frac{d\sigma_T}{d\Omega_K}(W, Q^2, \cos(\theta_K^*)). \quad (3.13)$$

Используя интерполированные значения структурной функции $\frac{d\sigma_T}{d\Omega_K}$, были восстановлены значения структурной функции $\frac{d\sigma_{TT}}{d\Omega_K}$. На рис. 3.5 приведены результаты данной процедуры. Процедура хорошо воспроизводит имеющиеся результаты, полученные из измерений на детекторе CLAS.

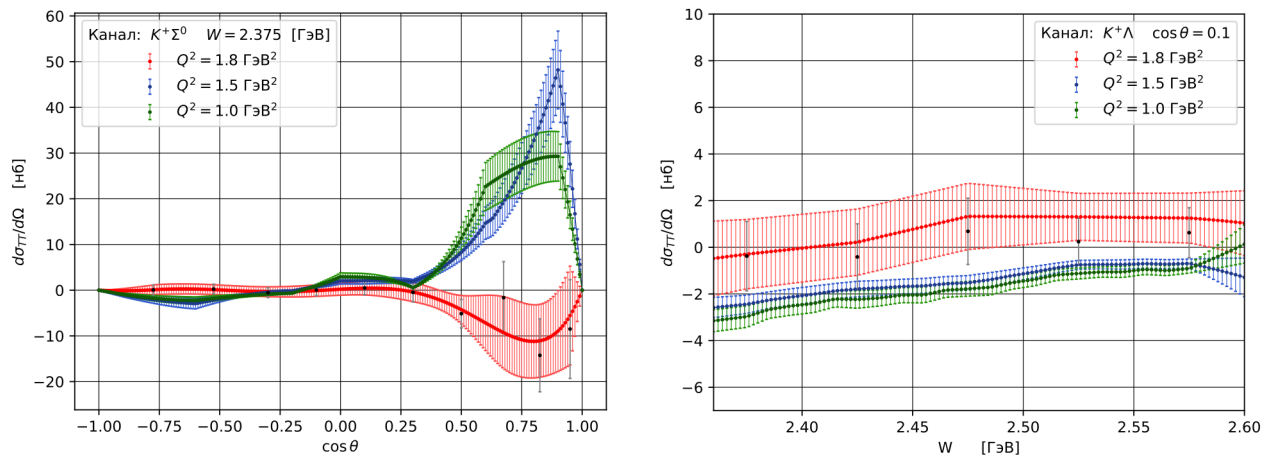


Рисунок 3.5 - Оценка структурной функции $\frac{d\sigma_{TT}}{d\Omega_K}$ в области

$2.17 < W < 2.65$ ГэВ и $Q^2 < 1.8$ ГэВ²: для каналов $K^+\Lambda$ (слева) и $K^+\Sigma^0$ (справа).

Экспериментальные данные (точки с неопределенностями) при $Q^2 = 1.8$ ГэВ² представлены в сравнении с результатами интерполяции (красная кривая). Синяя и зеленая кривые являются результатами интерполяции структурных функций $\frac{d\sigma_{TT}}{d\Omega_K}$ в $Q^2 = 1.5$ ГэВ², $Q^2 = 1.0$ ГэВ² и $\cos(\theta_K^*) = 0.1$ как зависимость от W (слева), и как зависимость от $\cos(\theta_K^*)$ при $W = 2.375$ ГэВ (справа).

3.5 Описание программы и веб-сайта для оценки сечений каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$

Приведенные процедуры интерполяции и экстраполяции структурных функций позволили рассчитать значения дифференциальных сечений при произвольных значениях $(W, Q^2, \cos(\theta_K^*))$ и энергии пучка. Дифференциальные сечения электророждения $K^+\Lambda$ и $K^+\Sigma^0$ рассчитываются согласно (3.1) и (3.5) из интерполированных/экстраполированных структурных функций. Сравнение с результатами из измеренных данных CLAS показано на рис. 3.6. Хорошее качество описания данных CLAS достигнуто во всей кинематической области $W < 2.65$ ГэВ и $Q^2 < 5.0$ ГэВ².

Экспериментальные данные по дифференциальным сечениям $K^+\Lambda$ и $K^+\Sigma^0$ электророждения представлены в виде интервалов по переменным $(W, Q^2, \cos(\theta_K^*))$. Для сравнения с результатами расчетов была разработана процедура усреднения рассчитанных дифференциальных сечений в пределах каждого интервала по $(W, Q^2, \cos(\theta_K^*))$ в измеренных данных. Процедура включает в себя:

1. Расчет структурных функций и дифференциальных сечений в каждом из интервалов по $(W, Q^2, \cos(\theta_K^*))$ на сетке с шагом по $\Delta W = 10$ МэВ, $\Delta Q^2 = 0.1$ ГэВ², $\Delta \cos(\theta_K^*) = 0.1$ и $\Delta \phi_\pi^* = 15^\circ$,
2. Усреднение полученных дифференциальных сечений по всем узлам сетки в пределах измеренных интервалов по $(W, Q^2, \cos(\theta_K^*))$:

$$\bar{\sigma}_N = \sum_{i=0}^N \frac{\sigma_i}{N} \xrightarrow{N \rightarrow \infty} \bar{\sigma},$$

где N – число узлов в пределах интервала по $(W, Q^2, \cos(\theta_K^*))$.

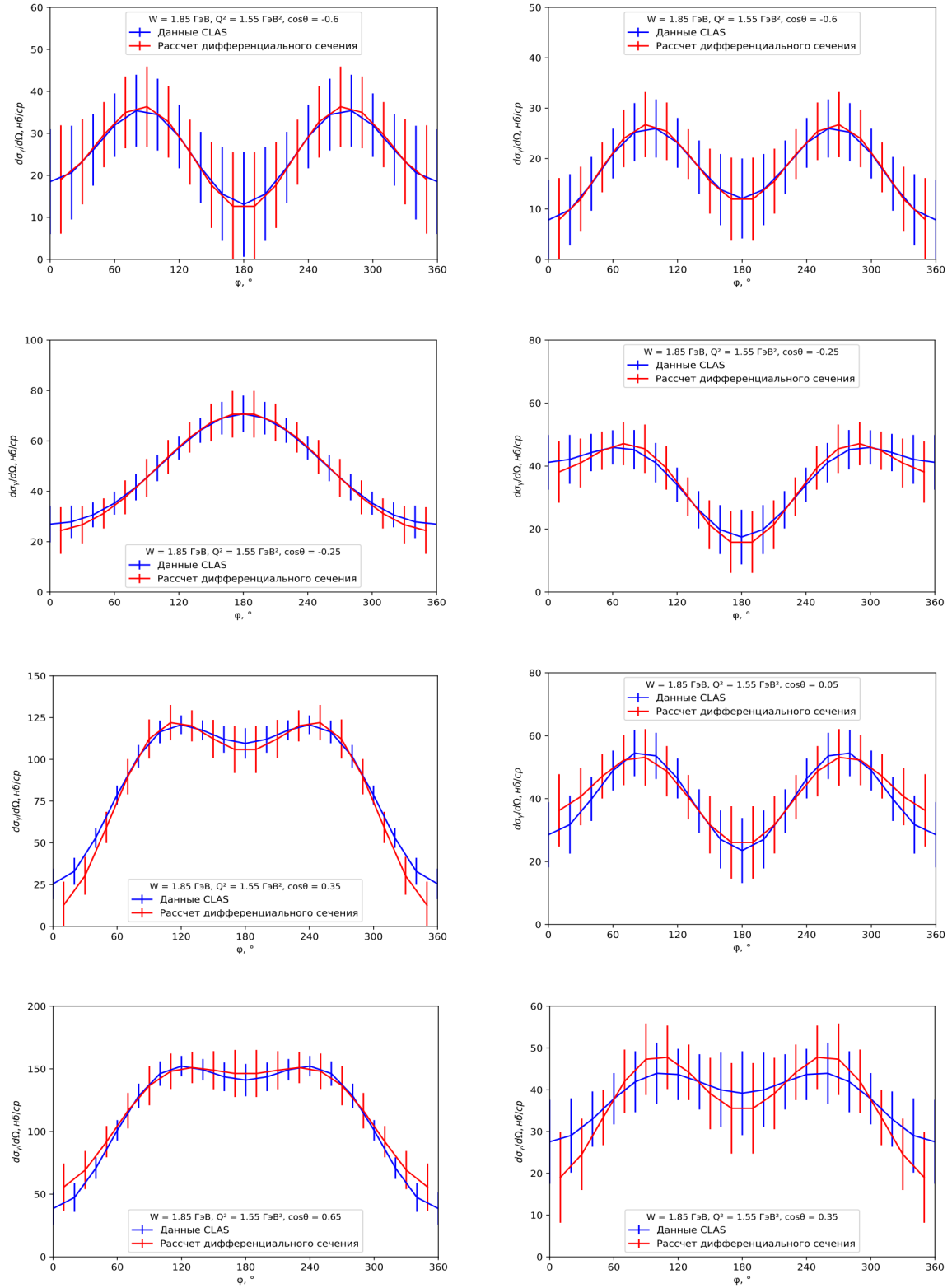


Рисунок 3.6 - Сравнение результатов прогнозирования дифференциального сечения $\frac{d\sigma_\gamma}{d\Omega_K}$ (красная линия) с экспериментальными данными (синяя линия) для канала $K^+ \Lambda$ (слева) и $K^+ \Sigma^0$ (справа) при различных значениях $W, Q^2, \cos(\theta_K^*)$.

Рис. 3.7 демонстрирует сравнение дифференциального сечения, полученного из экспериментальных данных по структурным функциям, с результатом описанной выше процедуры. Основываясь на представленных графиках, можно заключить, что полученные зависимости находятся в хорошем согласии.

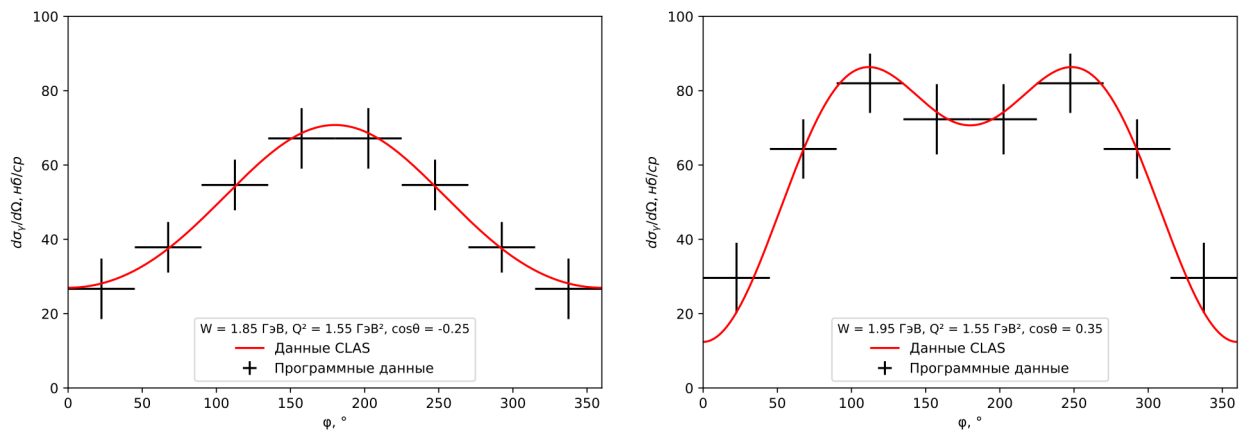


Рисунок 3.7 - Сравнение результатов усреднения рассчитанных сечений (черные точки) с оценкой дифференциального сечения на основе экспериментальных структурных функций (красная кривая) для канала $K^+\Lambda$ при различных значениях $W, Q^2, \cos(\theta_K^*)$.

3.5.1 Комплекс программ для оценки сечений

Программная реализация метода оценки дифференциальных сечений электророждения $K^+\Lambda$ и $K^+\Sigma^0$ представлена на двух языках: C++ и Python – с использованием различных библиотек для визуализации результата и интерполяции экспериментальных данных. Исходный код программы опубликован для общего доступа и может быть найден по ссылке [48].

Программа работает в трех режимах:

- оценка дифференциального сечения/структурных функций в заданной сетке значений $(E, W, Q^2, \cos(\theta_K^*), \phi_\pi^*)$;

- расчет средних значений дифференциальных сечений рассеяния в заданной ячейке фазового пространства;
- визуализация распределений дифференциального сечения/структурных функций по осям $(W, Q^2, \cos(\theta_K^*), \phi_\pi^*)$.

Комплекс программ также представлен на веб-сайте НИИ ядерной физики имени Д. В. Скобельцына МГУ [49].

Финальная версия программы включает в себя возможность расчета структурных функций и дифференциального сечения в любой точке фазового пространства реакции, среднего значения сечения в определенной ячейке и визуализацию распределений структурных функций, дифференциального сечения, а также средних значений в одномерных и двумерных представлениях.

4 Моделирование и анализ эксклюзивных каналов двухпионного фоторождения с применением алгоритмов машинного обучения

В данной главе описан основанный на генеративном моделировании подход к восстановлению детекторных эффектов в эксклюзивных реакциях с многочастичными конечными состояниями. Данное исследование демонстрирует сохранение корреляций между кинематическими переменными в многомерном фазовом пространстве. Проводится тестирование на синтетических данных фоторождения двух пионов, сгенерированных с использованием реалистичной модели в кинематике эксперимента CLAS в национальной исследовательской лаборатории имени Т. Джефферсона. Совмещение различных механизмов реакции, приводящих к одному и тому же конечному состоянию, а также сложные эффекты детектора CLAS представляют собой идеальный тест для анализа с применением искусственного интеллекта. Метод бутстрапа позволяет оценить систематическую неопределенность, связанную с применяемой процедурой. Тесты показывают, что генеративные состязательные сети (GAN) способны воспроизводить высоко коррелированные дифференциальные сечения даже при наличии детекторных искажений в обучающих данных, что является основой для применения этого метода к реальным экспериментальным данным.

Фоторождение двух пионов при энергиях фотонов в диапазоне нескольких ГэВ является важным процессом в адронной спектроскопии. Данный процесс широко используется для изучения ряда фундаментальных вопросов, таких как проблема «недостающих барионов», а также для демонстрации необходимости многочастичных конечных состояний для определения спектра неустойчивых возбужденных состояний. Хотя имеются обширные данные по фоторождению одиночного пиона, и соответствующая феноменология хорошо изучена, добавление третьей частицы в конечное состояние делает описание этой реакции значительно

более сложным. При фиксированной энергии фотона E_γ неполяризованное сечение фоторождения одного пиона описывается одной независимой переменной, тогда как для двух пионов требуются три дополнительных переменные. При энергиях пучка в несколько ГэВ наиболее статистически значимый набор данных был получен в эксперименте CLAS в национальной исследовательской лаборатории имени Т. Джефферсона [50]. Даже в этом случае некоторые ячейки в многомерном пространстве кинематических переменных остаются незаполненными или подвержены значительным статистическим флуктуациям, что влечет за собой существенные неопределенности в реконструкции механизмов реакции.

Ранее эту проблему решали путем анализа одной или двух переменных с последующим интегрированием по остальным. Однако во время интегрирования теряются корреляции между переменными, которые содержат важную физическую информацию, что делает результаты сильно зависимыми от модели. В этом контексте генеративные модели, основанные на машинном обучении, которые обучаются на исходных данных и создают новые синтетические данные, имитирующие оригинальное распределение, открывают новые возможности для извлечения физической информации при сохранении корреляций. Кроме того, данные модели позволяют выполнять анфолдинг (unfolding), то есть восстанавливать по экспериментальным данным распределения физических величин, очищенные от детекторных и фоновых искажений.

Применение машинного обучения для анфолдинга привлекло значительное внимание в последние годы в области физики частиц [51], и оно было применено в нескольких анализах [52, 53, 54]. Данная работа отличается от предыдущих исследований, таких как [51], явным использованием генеративных методов моделирования. Это исследование в значительной степени совпадает по тематике с работой [55].

Недавнее исследование показало возможность использования генеративных состязательных сетей для анфолдинга на уровне событий в инклюзивных реакциях электророждения [56]. Анализ позволил точно восстановить сечения по одной

переменной. В данной работе эта методика расширяется на многочастичные конечные состояния, впервые демонстрируя, что предложенный метод GAN может быть использован для воспроизведения процессов рассеяния в фазовом пространстве высокой размерности. В частности, оптимизируется модель машинного обучения для фоторождения двух пионов в кинематике CLAS.

Данное исследование служит хорошим тестовым полигоном для оценки эффективности анализа с использованием методов машинного обучения. Наличие барионных и мезонных резонансов с различными механизмами рождения, перекрывающихся в ограниченном фазовом пространстве, порождает сложные структуры и корреляции. Кроме того, высокая неоднородность отклика детектора CLAS вводит дополнительные сложности и искажения, усложняя анализ. Для тестирования и валидации предложенного подхода генерируются Монте-Карло (MC) синтетические данные с использованием реалистичной модели фоторождения двух пионов. Синтетическая копия создается с помощью анфолдинг-GAN, обученной на псевдоданных, включающих детекторные эффекты, моделируемые GEANT [57]. Такой подход эквивалентен обучению GAN на экспериментальных данных. Анфолдинг детекторных эффектов производится с помощью «детекторно-моделированной» GAN (обученной на данных, моделирующих отклик детектора), независимо обученной на втором наборе псевдоданных, полученных методом Монте-Карло, сгенерированных согласно фазовому пространству и пропущенных через модель детектора GEANT. Качество метода оценивается количественным сравнением сгенерированных MC данных и их синтетической копии. Данный тест, основанный на MC-псевдоданных, является необходимым шагом перед применением предложенной методики к экспериментальным данным.

4.1 Двухпионное фоторождение

4.1.1 Описание реакции двухпионного фоторождения на протоне

Конечное состояние $\pi\pi N$ вносит существенный вклад в суммарное сечение фоторождения на протонах при энергиях центра масс $W \lesssim 2.5$ ГэВ. Исследования этого конечного состояния позволили существенно расширить знания о спектре возбужденных состояний нуклона (N^*) и их амплитудах фотовозбуждения. Квантовые числа этих резонансов можно определить, изучая корреляции между инвариантной массой и угловыми зависимостями их продуктов распада. Теоретические оценки, основанные на феноменологических подходах [58, 59, 60, 61, 62], методах, основанных на формализме Швингера [61, 62], а также вычислениях на решетке КХД [63], предсказывают больше состояний, чем наблюдается в экспериментах (см. обзоры [64, 65, 66, 67]), что известно как проблема «недостающих барионов».

Одним из способов повысить чувствительность к наиболее труднодоступным состояниям является наложение условий согласованности путем одновременного анализа нескольких конечных состояний, где $\pi\pi N$ играет ключевую роль для резонансов с массами выше 1.6 ГэВ. Такой подход позволяет выделить и устранить вклад коррелированных нерезонансных процессов и практически независимо от модели извлечь параметры резонансов. Кроме того, объединение данных по фоторождению и электророждению позволило недавно идентифицировать перекрывающиеся резонансы с одинаковыми квантовыми числами, например, $N(1720)$ и $N'(1720)$ [68, 46].

В указанной реакции, по распределению инвариантной массы пары $\pi\pi$, можно исследовать мезонные резонансы, такие как ρ или $f_2(1270)$ [69]. Хотя свойства этих резонансов хорошо изучены, механизмы их фоторождения до сих пор остаются недостаточно понятными. В области низких энергий $W \lesssim 2$ ГэВ можно

исследовать вклад отдельных N^* состояний в процессы рождения мезонов. В более высокоэнергетической области, выше зоны резонансов N^* , реакция хорошо описывается в рамках теории Редже (Regge theory) [70]. Непрерывный характер перехода между указанными энергетическими режимами обуславливает аналитические трудности в исследовании промежуточной области. Существует исследование, в котором предложена формализация для описания процесса рождения одиночных π или η мезонов [71]. Расширение этого подхода на дважды псевдоскалярные конечные состояния требует учета полной многомерной кинематической зависимости. В частности, перед исследованием более сложных каналов $\eta\pi N$ и $\eta'\pi N$, в которых, согласно теоретическим ожиданиям, могут проявляться экзотические адроны, необходимо детально изучить механизмы рождения мезонов в конечном состоянии $\pi\pi N$.

4.1.2 Кинематика $\gamma p \rightarrow \pi^+ \pi^- p$

Измерение трехчастичного конечного состояния в процессе фоторождения двух пионов представляет значительную экспериментальную задачу. Недавно стало доступно большое количество данных по фоторождению $\pi^+ \pi^- p$ благодаря измерениям, проведенным коллаборацией CLAS при энергиях $W \leq 2.9$ ГэВ [65, 72]. Для данной энергии столкновения дифференциальное сечение этого процесса зависит от пяти независимых переменных, которые могут быть выбраны следующим образом: инвариантные массы системы двух пионов, $M_{\pi^+ \pi^-}$, и системы протон-пион, $M_{p\pi^-}$, а также три угловые переменные в системе центра масс. Два угла — это полярный угол θ_{π^+} (с осью z , направленной вдоль импульса фотона) и угол $\alpha_{[\pi^+ p][\pi^- p']}$ между плоскостью, содержащей исходный протон-мишень p и импульс π^+ , и плоскостью, содержащей импульсы π^- и рассеянного протона p' . Альтернативный выбор — это замена θ_{π^+} на переданный инвариантный момент t_{π^+} , который определяется как квадрат разности четырехимпульсов фотона и π^+ . Пятая

переменная ϕ — азимутальный угол π^- относительно плоскости, содержащей импульс фотона и вектор его поляризации. Данный угол имеет значение только в экспериментах с поляризованным пучком или поляризованной мишенью. Для неполяризованных данных можно определить ϕ , направляя вектор поляризации в произвольное направление, что приводит к сечению, не зависящему от ϕ . Возможны альтернативные наборы переменных, например, $M_{p\pi^+}$ (инвариантная масса системы протон- π^+), t_{π^-} (переданный момент между фотоном и π^-), t (переданный момент между протоном-мишенью и рассеянным протоном), или $\cos\theta$ (косинус угла между протоном-мишенью и рассеянным протоном в системе центра масс).

Многомерные анализы становятся стандартом в современных экспериментах с высокой статистикой, хотя они остаются вычислительно сложными [54]. Однако в некоторых реакциях могут возникнуть ограничения, связанные с недостаточной статистикой. В частности, прямая реконструкция событий фоторождения $\pi^+\pi^-p$ при заданном значении W на $5D$ -сетке (или $4D$, если интегрировать по углу ϕ) с приемлемым размером ячейки для физического анализа оказывается затруднительной задачей. Даже самый статистически насыщенный набор данных по фоторождению $\pi^+\pi^-p$, собранный в CLAS [72], приводит к небольшому числу событий в $4D$ -ячейках (обычно менее 10 событий на ячейку). В работе теоретические кривые подгонялись под одномерные проекции, полученные путем интегрирования скорректированного на эффективность детектора $5D$ -распределения по оставшимся четырем переменным. Однако такой метод нивелирует корреляции, имеющиеся в исходных данных, что приводит к потере значительной части информации, содержащейся в совместных распределениях.

В данной работе предпринимается попытка решения указанной проблемы посредством применения методов машинного обучения. Для иллюстрации этого на рис. 4.1 представлены примеры двумерных распределений и их одномерных

проекций, измеренных в эксперименте CLAS без учета эффективности детектора [46].

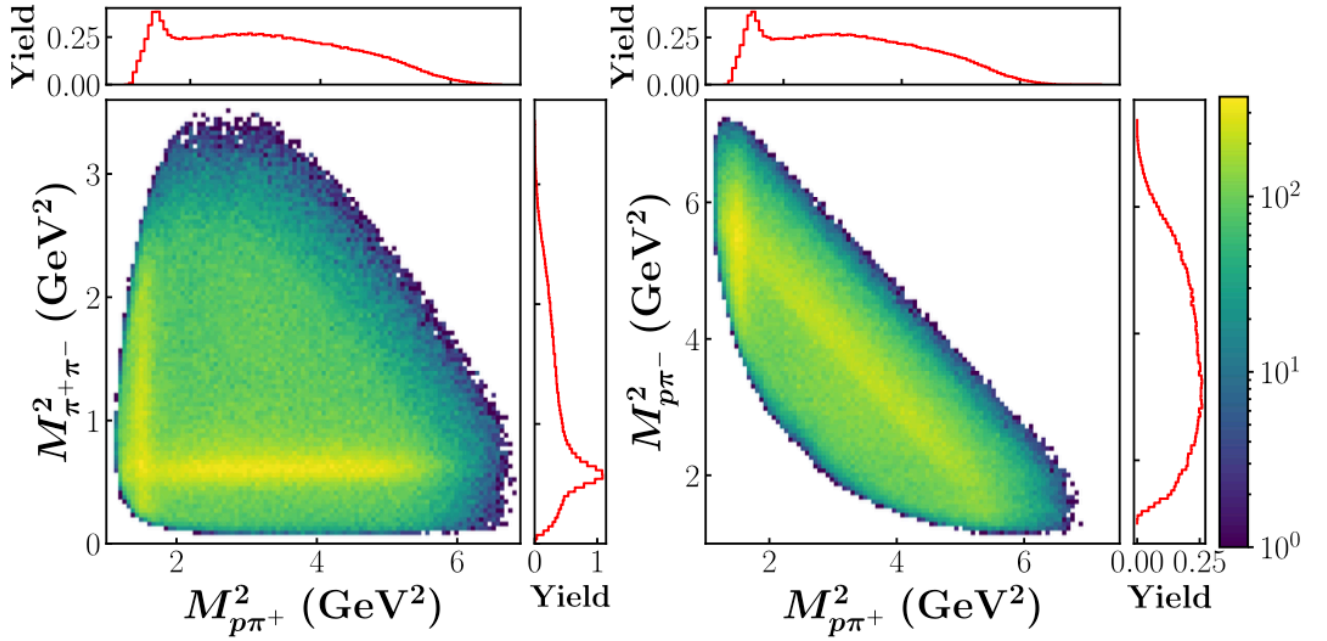


Рисунок 4.1 — Примеры нормализованных 2D-распределений выходов и их 1D-проекций, измеренных в эксперименте CLAS (до анфолдинга детекторных эффектов). Графики показывают распределения инвариантной массы для выходов $M_{\pi^+\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (слева) и $M_{p\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (справа). Для каждого 2D-графика соответствующие 1D-проекции показаны сверху и справа. Распределения смещены за счет значительных вкладов резонансов Δ^{++} и ρ , которые проявляются в виде полос в 2D-распределениях и пиков в соответствующих 1D-проекциях. В частности, Δ^{++} появляется как вертикальная полоса на 2D-графиках и как пик в 1D-проекции $M_{p\pi^+}^2$. Резонанс ρ проявляется в виде горизонтальной полосы на левом 2D-графике, диагональной полосы на правом 2D-графике и пика в 1D-распределении $M_{\pi^+\pi^-}^2$. Из правых графиков видно, что при выборе $M_{p\pi^+}^2$ и $M_{p\pi^-}^2$ в качестве независимых переменных существование ρ теряется, если анализировать только 1D-проекции.

Эти распределения сразу же показывают наличие промежуточных резонансов, которые проявляются как усиления в инвариантных массах систем, в которых они распадаются. Например, полоса при $M_{p\pi^+}^2 \approx 1.5 \text{ ГэВ}^2$ соответствует барионному резонансу $\Delta(1232)$, который появляется как промежуточное неустойчивое состояние в реакции $\gamma p \rightarrow \Delta^{++}\pi^- \rightarrow \pi^+\pi^-p$. Полоса, центрированная при $M_{\pi^+\pi^-}^2 \approx$

0.6 ГэВ^2 , соответствует мезонному резонансу $\rho(770)$ в реакции $\gamma p \rightarrow p \rho^0 \rightarrow \pi^+ \pi^- p$. Оба резонанса проявляются в виде пиков в соответствующих одномерных проекциях. Однако при рассмотрении только одномерных проекций резонанс может остаться незамеченным, если соответствующее распределение инвариантной массы не анализируется; например, в проекции $M_{p\pi^-}^2$ пик отсутствует. Это пример потери информации, связанной с корреляциями. Кроме того, из-за квантовой интерференции процессы рождения ρ^0 и Δ^{++} не являются независимыми, и невозможно однозначно приписать конкретное событие одному из этих процессов. Эта интерференция проявляется в корреляциях между инвариантными массами, и она может быть частично утрачена при использовании одномерных проекций.

4.1.3 Двухпионное фоторождение на данных с детектора CLAS

Спектрометр CLAS лаборатории имени Т. Джефферсона основан на тороидальном магнитном поле с максимальной напряженностью около 1.25 Т, которое изгибает траектории заряженных частиц, образующихся в результате адронного взаимодействия, вдоль полярных углов θ_{lab} (где ось z направлена вдоль направления движения пучка фотонов), при этом сохраняя их азимутальные углы ϕ_{lab} . Полярность поля определяла, будут ли положительно или отрицательно заряженные частицы изгибаться к или от оси пучка в чувствительной области детектора. Система из трех слоев многопроволочных дрейфовых камер [73] обеспечивала информацию об импульсе с разрешением σ_p/p , варьирующимся от 0.5% до 1.0% в зависимости от кинематики. Идентификация заряженных адронов осуществлялась с помощью времяпролетных сцинтилляционных детекторов. Эксперименты по фоторождению проводились с тормозным фотонным пучком, полученным путем рассеяния непрерывного электронного пучка с длиной волны в 8×10^{-5} от толщины фольги. Тормозная система идентификации [74] с

разрешением энергии фотона 0.1% использовалась для измерения энергии фотона в каждом зарегистрированном событии. Таргетная ячейка представляла собой цилиндр из майлара диаметром 4 см и длиной 40 см, заполненный жидким водородом при температуре 20.4 К.

Представленные экспериментальные условия соответствуют эксперименту *g11*, проведенному на CLAS в 2004 году. Во время эксперимента тороидальное поле было настроено таким образом, что положительно заряженные частицы изгибались от оси пучка. Геометрическое разрешение детектора для каждой положительно заряженной частицы в соответствующем кинематическом диапазоне составляло около 40% и несколько меньше для отрицательно заряженных частиц, которые изгибались к оси пучка и, как следствие, выходили за пределы разрешения детектора. Энергия первичного электронного пучка составляла 4.02 ГэВ, обеспечивая идентификацию фотонного пучка в диапазоне энергий от 0.8 до 3.8 ГэВ. В данном анализе основное внимание уделяется области самых высоких энергий, 3.0 – 3.8 ГэВ, которая была проанализирована в [72].

Эксклюзивная реакция $\gamma p \rightarrow \pi^+ \pi^- p$ была выделена путем регистрации протона и π^+ в спектрометре CLAS, в то время как π^- реконструировался на основе четырехимпульсов зарегистрированных частиц с использованием метода недостающей массы. Таким образом, обеспечивалась эксклюзивность реакции, сводя к минимуму вклад процессов от многопионного фона. В анализе использовались только события в пределах определенного объема фазового пространства, чтобы исключить области на границах разрешения детектора. Были установлены ограничения на минимальный импульс протона, а также минимальный и максимальный полярный угол адрона. После всех ограничений было идентифицировано приблизительно 40 миллионов событий, произведенных в процессе эксклюзивного фоторождения двух пионов, что делает набор данных *g11* крупнейшим по объему событий для этой реакции в указанном диапазоне энергий фотона. Подробности анализа эксперимента *g11* можно найти в [72].

4.2 Моделирование методом Монте-Карло

Псевдоданные, соответствующие фоторождению двух пионов в кинематике CLAS, были сгенерированы с использованием двух различных генераторов событий Монте-Карло. Один из них представляет собой чистое фазовое пространство, а другой использует реалистичную физическую модель. Сгенерированные псевдоданные затем пропускались через модель детектора GEANT для учета его влияния на экспериментальные данные. Модель GEANT включала геометрию детектора, разрешение и эффективность регистрации, а также влияние неоднородности разрешения детектора.

Эффекты детектора оценивались с помощью первого генератора Монте-Карло, основанного на распределении «чистого» (неизмененного фоновыми и детекторными процессами) фазового пространства. Для выполнения теста согласованности (closure test) был использован второй генератор Монте-Карло, включающий реалистичную физическую модель. Использование двух различных генераторов минимизирует модельную зависимость при выделении исходной информации и имитирует реальную ситуацию, когда эффекты детектора оцениваются с помощью симуляций, которые похожи, но не идентичны экспериментальным распределениям.

4.2.1 Генераторы событий фоторождения двух пионов

Два генератора событий Монте-Карло моделируют взаимодействие падающего неполяризованного фотонного пучка с тормозным спектром в диапазоне энергий 3.0 – 3.8 ГэВ с неподвижным протоном-мишенью. При выбранных переменных распределения событий пропорциональны дифференциальному сечению, а следовательно, квадрату амплитуды процесса A , суммированной по поляризациям:

$$\begin{aligned}
& \frac{d^5\sigma}{dM_{p\pi^-}^2 dM_{\pi^+\pi^-}^2 dt_{\pi^+} d\alpha_{[\pi^+p][\pi^-\pi']} d\phi} \\
& \propto \left[\left(W^2 - (M_{p\pi^-} + m_\pi)^2 \right) \left(W^2 - (M_{p\pi^-} - m_\pi)^2 \right) \right]^{-1/2} \\
& \times \sum_{pol} \left| A(M_{p\pi^+}^2, M_{\pi^+\pi^-}^2, \cos\theta_{\pi^-}, \alpha_{[\pi^+p][\pi^-\pi']}) \right|^2.
\end{aligned} \tag{4.1}$$

Первый генератор Монте-Карло, называемый PS-МС (фазовое пространство Монте-Карло), распределяет финальные состояния событий в соответствии с фазовым пространством системы $\pi^+\pi^-p$. Это эквивалентно предположению о константной амплитуде процесса. Хотя это явно нереалистично, поскольку фоторождение двух пионов имеет гораздо более сложную структуру, такой подход позволяет равномерно распределять события по всему пространству в кинематике реакции. Одномерные проекции распределений событий, сгенерированных с помощью PS-МС, показаны на рис. 4.2, а двумерные распределения — на рис. 4.3.

Второй генератор событий Монте-Карло, который называется RE-МС (реалистичный Монте-Карло), учитывает квадрат амплитуды как некогерентную сумму трех доминирующих промежуточных резонансов, наблюдаемых в реакции $\gamma p \rightarrow (p\rho^0, \Delta^{++}\pi^-, \Delta^0\pi^+) \rightarrow \pi^+\pi^-p$, с добавлением приблизительно 10% постоянного вклада, имитирующего нерезонансное фоторождение двух пионов. Каждый процесс взвешен в соответствии с его вкладом в полное сечение, как это представлено в [75]. Угловые распределения, относящиеся к рождению резонансов, параметризованы на основе измеренных дифференциальных сечений из совпадающей базы данных. Распады $\rho \rightarrow \pi\pi$ и $\Delta \rightarrow p\pi$ описаны с учетом правильной спиновой структуры, используя элементы матрицы распада. Одномерные и двумерные проекции событий, сгенерированных с помощью RE-МС, показаны на рис. 4.4 и рис. 4.5. Данный подход не учитывает интерференционные члены между промежуточными резонансами. Тем не менее, полученные распределения обеспечивают разумное описание экспериментальных данных, демонстрируя резонансные структуры в инвариантных массах и правильное угловое распределение частиц в финальном состоянии.

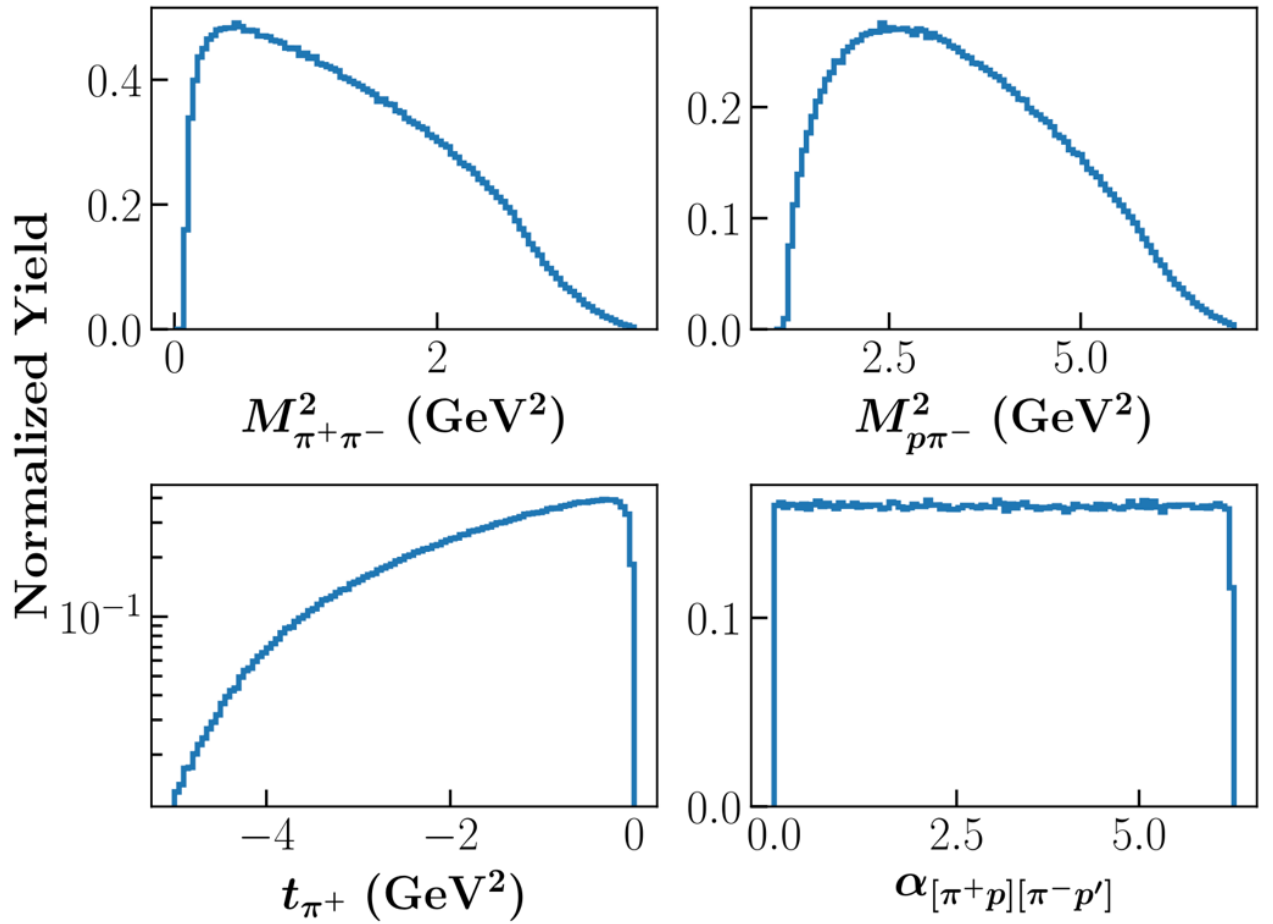


Рисунок 4.2 - Нормализованные 1D-проекции событий, сгенерированных с использованием PS-MC: квадрат инвариантной массы $M_{\pi^+\pi^-}^2$ (вверху слева), квадрат инвариантной массы $M_{p\pi^-}^2$ (вверху справа), квадрат четырехимпульса, переданного от мишени протону отдачи t_{π^+} (внизу слева), а также угол $\alpha_{[\pi^+p][\pi^-p']}$ в системе центра масс (CM) между плоскостью, содержащей начальный импульс мишени p и импульс π^- , и плоскостью, образованной импульсами π^+ и рассеянного протона (внизу справа).

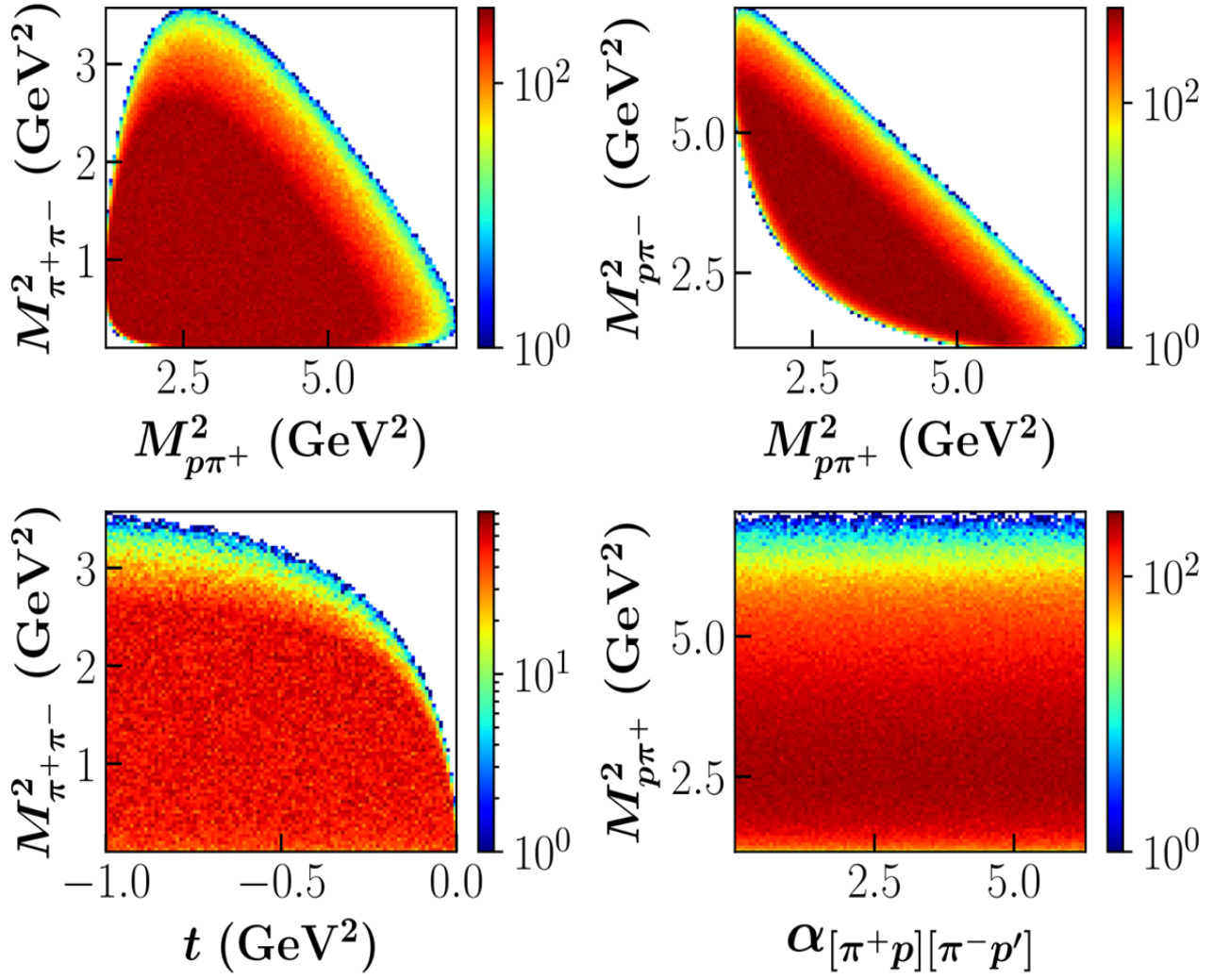


Рисунок 4.3 - Нормализованные 2D-распределения, сгенерированные с использованием PS-МС: распределения инвариантной массы $M_{\pi^+\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (вверху слева) и $M_{p\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (вверху справа), инвариантная масса $M_{\pi^+\pi^-}^2$ по сравнению с переданным импульсом t (внизу слева), а также угол в системе центра масс (СМ) по сравнению с $M_{p\pi^+}^2$ (внизу справа).

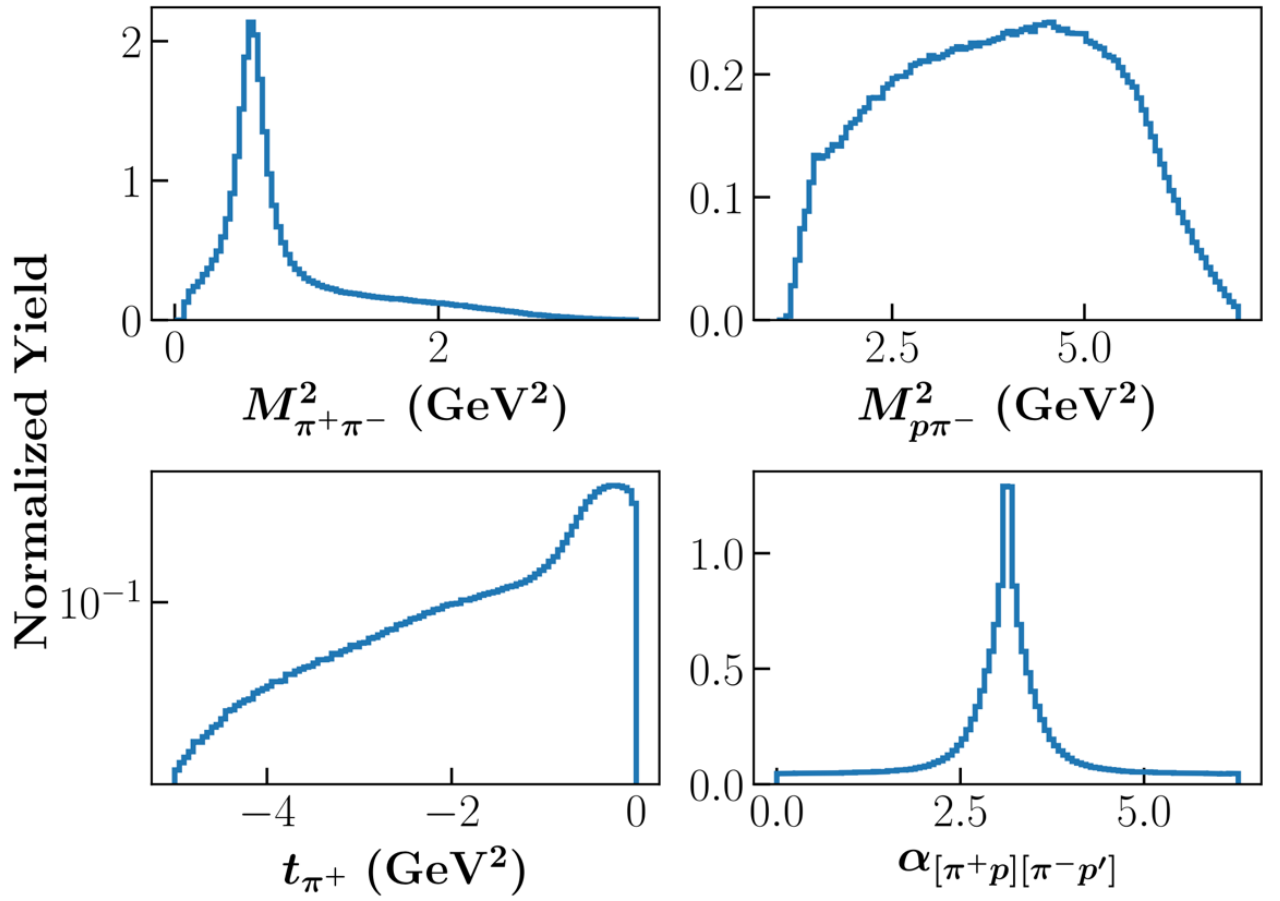


Рисунок 4.4 - Нормализованные 1D-проекции событий, сгенерированных с использованием RE-МС: квадрат инвариантной массы $M^2_{\pi^+\pi^-}$ (вверху слева), квадрат инвариантной массы $M^2_{p\pi^-}$ (вверху справа), квадрат четырехимпульса, переданного от мишени откатившемуся протону t_{π^+} (внизу слева), а также угол $\alpha_{[\pi^+p][\pi^-p']}$ в системе центра масс (CM) между плоскостью, содержащей начальный импульс мишени p и импульс π^- , и плоскостью, образованной импульсами π^+ и рассеянного протона (внизу справа).

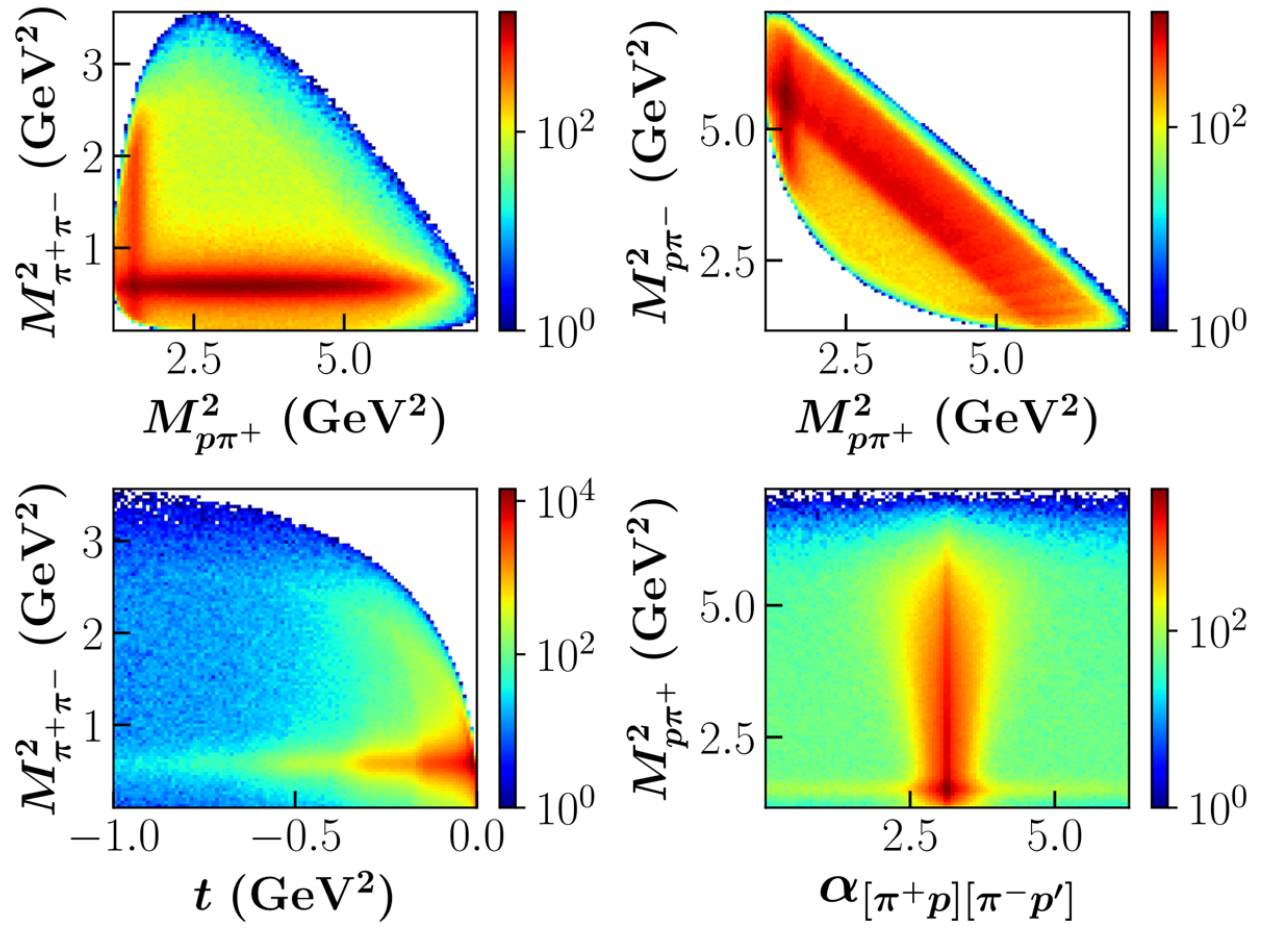


Рисунок 4.5 - Нормализованные 2D-распределения, сгенерированные с использованием RE-МС: распределения инвариантной массы $M_{\pi^+\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (вверху слева) и $M_{p\pi^-}^2$ по сравнению с $M_{p\pi^+}^2$ (вверху справа), инвариантная масса $M_{\pi^+\pi^-}^2$ по сравнению с переданным импульсом t (внизу слева), а также угол в системе центра масс (CM) по сравнению с $M_{p\pi^+}^2$ (внизу справа).

4.2.2 Моделирование детектора CLAS

Отклик детектора CLAS был смоделирован с использованием стандартного пакета моделирования на основе GEANT, GSIM, который используется и разрабатывается в коллаборации CLAS [76]. GSIM включает центральный управляющий модуль, который вызывает несколько независимых пакетов, включающих описание геометрии и отклика различных компонентов детектора. Постобработка данных выполнялась с помощью кода GSIM-Post-Processor (GPP), который корректировал выходные данные GSIM, чтобы лучше соответствовать экспериментальным характеристикам детектора, включая статистические флуктуации и мертвые зоны.

Реконструкция событий осуществлялась с использованием программного обеспечения RECSIS [77], применяемого для обработки экспериментальных данных. В данной работе термины «REC» или «детекторный уровень» обозначают набор псевдоданных после обработки в детекторе, а «GEN» или «уровень вершины» - истинные события, сгенерированные генератором Монте-Карло до учета эффектов детектора.

Разрешение детектора CLAS не является равномерным. Оно ограничено по азимутальному углу ϕ_{lab} за счет наличия шести магнитных катушек тороида, а также по полярному углу θ_{lab} из-за ограниченного покрытия дрейфовыми камерами, калориметром и времяпролетными датчиками. Дополнительное ограничение касается минимального принимаемого импульса заряженных адронов, который обусловлен потерями энергии в материалах на пути частицы, а также эффектом тороидального магнитного поля, изгибающего низкоэнергетические частицы за пределы разрешения детектора. Ограниченное разрешение CLAS приводит к уменьшению числа событий REC по сравнению с GEN, поскольку не все сгенерированные события реконструируются. Влияние разрешения CLAS на переменные в лабораторной системе координат показано на рис. 4.6.

Кроме того, детектор CLAS имеет конечное разрешение, которое приводит к "размытию" измеряемых кинематических переменных, в результате чего наблюдаются расхождения между REC и GEN даже в тех случаях, когда событие зарегистрировано. Размытие влияет на реконструируемые импульсы любых детектируемых частиц в пределах разрешения CLAS, создавая искажения, зависящие от их четырехимпульсов. Рис. 4.7 демонстрирует разрешение для импульса и полярного угла π^+ в зависимости от истинного значения переменной, соответствующей уровню GEN, а также одномерные проекции, соответствующие относительному разрешению CLAS. Подгонка этих данных двумя нормальными распределениями дала значения $\delta p/p \approx 0.8\%$ и $\delta \theta/\theta \approx 0.5\%$. Аналогичные эффекты наблюдаются и для других частиц.

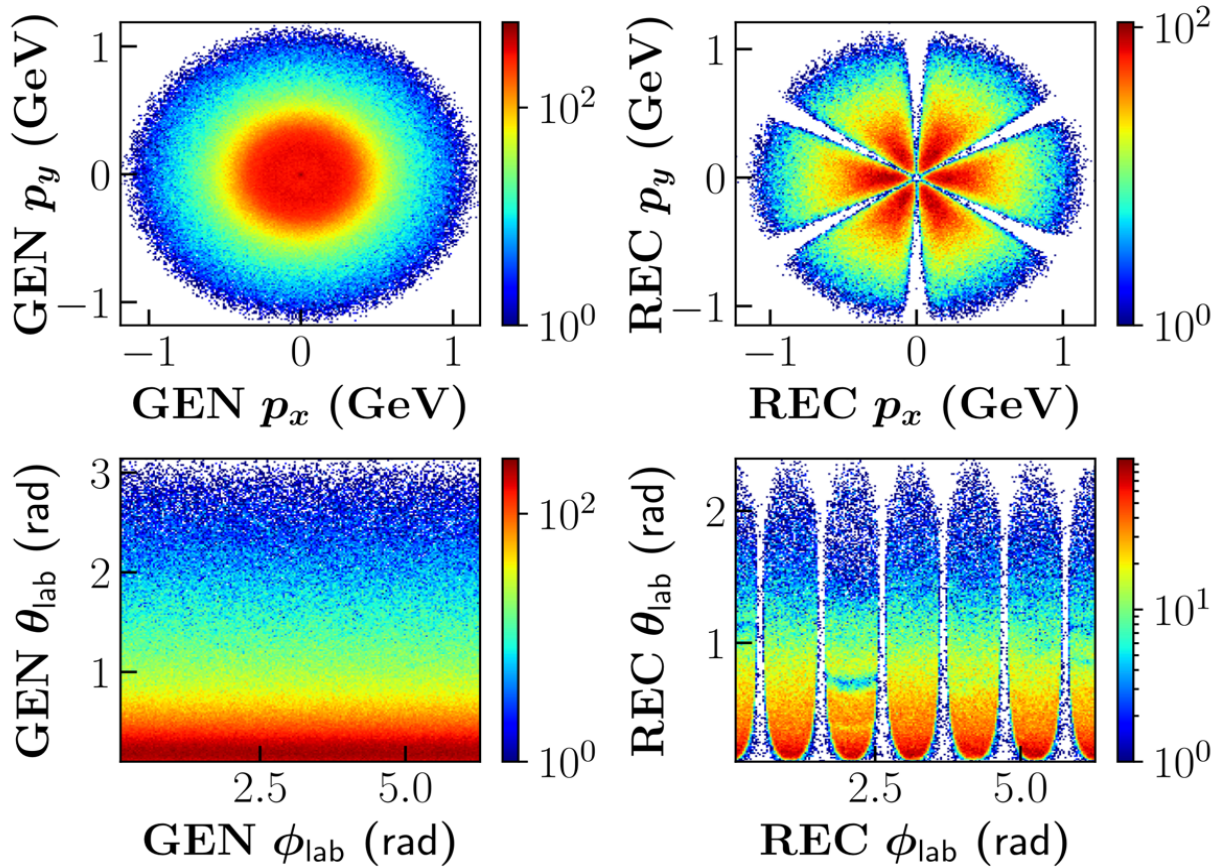


Рисунок 4.6 - Кинематические переменные π^+ в лабораторной системе отсчета, сгенерированные (GEN) с помощью RE-МC (слева) и реконструированные (REC) детектором CLAS (справа): p_x по сравнению с p_y (верхние графики) и θ_{lab} по сравнению с ϕ_{lab} (нижние графики).

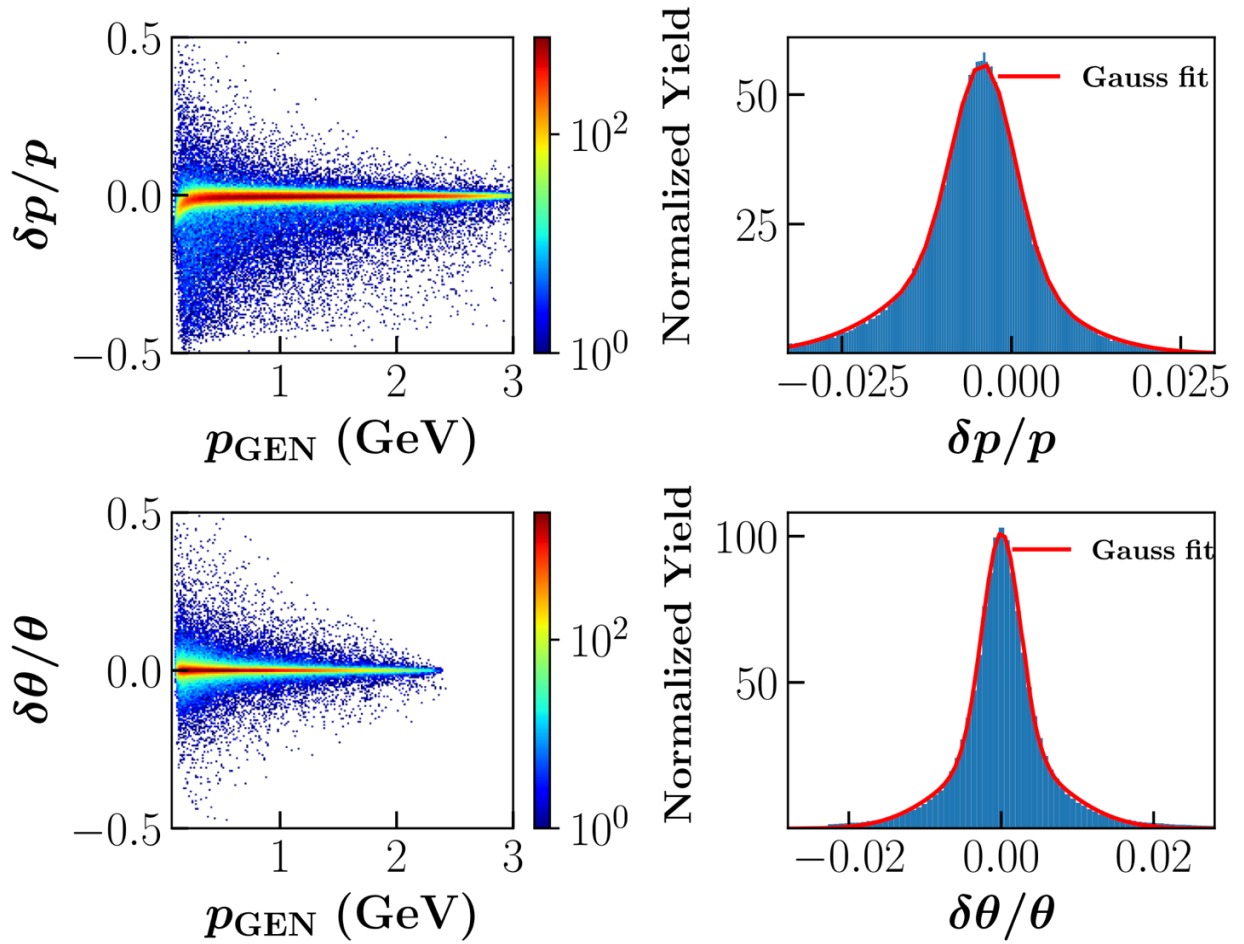


Рисунок 4.7 - Размывание кинематических переменных π^+ в лабораторной системе отсчета. Верхняя панель: относительное разрешение по импульсу $(p_{REC} - p_{GEN})/p_{GEN}$ в зависимости от p_{GEN} . Нижние графики: относительное угловое разрешение, определенное как $(\theta_{REC} - \theta_{GEN})/\theta_{GEN}$, в зависимости от p_{GEN} . На правых графиках представлены 1D-проекции 2D-распределений в виде гистограмм, а красные линии обозначают аппроксимацию двойным гауссовским распределением. Псевдоданные события были сгенерированы с помощью RE-MC.

Разрешение детектора CLAS достаточно высокое, что позволяет использовать метод недостающей массы для идентификации эксклюзивной реакции на фоне многопионных событий. Данный метод использует известные параметры начального состояния и зарегистрированных частиц для вычисления инвариантной массы недостающей системы, соблюдая законы сохранения энергии и импульса. Если все частицы обнаружены, недостающая масса равна нулю. Если одна частица отсутствует, то в распределении недостающей массы наблюдается пик, который соответствует ее массе. Если теряются две или более частиц,

распределение не имеет пика и становится гладким. Данный метод применим только в том случае, если разрешение эксперимента достаточно для того, чтобы отделить пик недостающей массы от многопионного фона. Очевидно, что чем больше частиц детектируется, тем хуже становится разрешение метода недостающей массы из-за накопления ошибок при вычислении. Когда недостающая частица идентифицирована, ее четырехимпульс определяется из условий сохранения энергии и импульса, и конечное состояние можно полностью реконструировать.

В фоторождении двух пионов ограничение одной недостающей частицы соответствует следующим топологиям (недостающая частица указана в скобках): $p\pi^+(\pi^-)$, $p\pi^-(\pi^+)$, $\pi^+\pi^-(p)$ и $\pi^+\pi^-p$ (все три частицы детектируются). С учетом разрешения CLAS, распределение событий между различными топологиями значительно варьируется, с отношением (100: 37: 30: 35) для указанных топологий соответственно. В данной работе анализируется топология $p\pi^+(\pi^-)$, поскольку данная топология в REC дает наибольшее количество данных, однако аналогичные выводы применимы и к другим топологиям. Каждая из топологий соответствует различным областям фазового пространства, и их комбинация расширяет кинематический охват измерений, смягчая эффект ограниченного разрешения детектора.

На рис. 4.8 представлено распределение недостающей массы для топологии $p\pi^+(\pi^-)$. Эксклюзивное финальное состояние выделяется выбором событий из области пика. В данном моделировании рассмотрены только двухпионные финальные состояния, поэтому многопионный фон отсутствует. В реальном эксперименте $g11$ значительный многопионный фон заполняет положительную часть спектра отсутствующей массы, но он удаляется в процессе анализа для выделения эксклюзивной реакции.

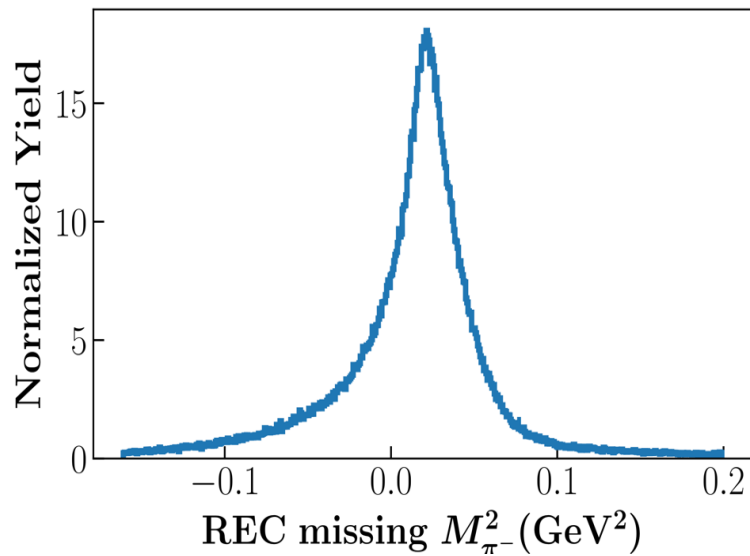


Рисунок 4.8 - Спектр недостающей массы π^- в топологии $p\pi^+(\pi^-)$ на основе REC-данных.

Псевдоданные для событий сгенерированы на основе RE-МС.

4.3 Методология анфолдинга на основе GAN

Генеративные состязательные сети (GAN), являющиеся разновидностью нейронных сетей, привлекли значительное внимание в последние годы. Данный алгоритм является основой для мощных генеративных моделей, высокоэффективных при создании синтетических данных, находящих применение в различных областях науки и культуры [78]. Архитектура типичной GAN включает сеть-генератор, которая учится генерировать данные, и сеть-дискриминатор, которая учится различать сгенерированные псевдоданные и реальные (эталонные) данные. Данные сети обучаются поочередно в конкурентной среде, где генератор пытается создавать более реалистичные данные, чтобы "обмануть" дискриминатор, а дискриминатор старается правильно идентифицировать сгенерированные данные. Данный итеративный процесс приводит к все более реалистичным результатам, приближаясь к цели — созданию синтетических данных, неотличимых от эталонных.

Сети GAN нашли широкое применение во многих сферах, таких как синтез изображений, генерация текста, композиция музыки и видео, и показали

впечатляющие результаты [79]. В области синтеза изображений GAN используются для создания сверхреалистичных изображений, визуально неотличимых от реальных, что имеет многочисленные практические приложения. Однако успешное обучение сетей GAN может быть чрезвычайно сложной задачей. Многие модели GAN сталкиваются со значительными проблемами, такими как коллапсирование модели, отсутствие сходимости, колебания параметров модели, дестабилизация, исчезающие градиенты и переобучение, что приводит к несбалансированному обучению генератора и дискриминатора [80]. В отличие от типичных приложений сетей GAN, успех генеративного моделирования событий в ядерной физике и физике элементарных частиц зависит от способности точно воспроизводить корреляции между импульсами частиц, что становится все более сложной задачей при увеличении размерности. Кроме того, многомерные распределения импульсов событий, связанных с реакциями в ядерной физике, такими как фоторождение двух пионов, рассматриваемое в данной работе, демонстрируют крайне сложные закономерности. Разработка соответствующей архитектуры GAN, способной одновременно воспроизводить все корреляции между импульсами частиц и воспроизводить многомерные распределения, представляет собой существенную исследовательскую и инженерную задачу.

Генеративные модели машинного обучения становятся все более популярными в качестве эффективных инструментов быстрого моделирования в различных научных областях, включая физику высоких энергий и ядерную физику [81]. В отличие от традиционных методов моделирования, которые основываются на теоретической модели для описания реакции, генеративные модели обучаются на больших наборах данных (без использования модельных предположений и других априорных знаний о системе) и используют полученные знания для генерации новых событий с высокой точностью. Сети GAN зарекомендовали себя как мощный инструмент в области быстрого моделирования, который позволяет генерировать события, которые максимально приближены к эталонным данным, воспроизводя основные физические процессы и их распределения [82].

Кроме того, сети GAN применялись для решения задачи моделирования детекторных эффектов при быстром моделировании [55, 56]. Это применение GAN позволяет уменьшить разрыв между моделированными и экспериментальными данными, обеспечивая более точное и реалистичное моделирование для анализа экспериментов. Исчерпывающий обзор существующих генераторов событий на основе машинного обучения можно найти в [83]. В данном исследовании был применен архитектурный подход наименьших квадратов для GAN, который включает замену функции потерь кросс-энтропии в дискриминаторе стандартной GAN на сумму квадратов ошибок.

Далее дано описание GAN, используемой для генерации синтетических данных, воспроизводящих RE-МС псевдоданные реакции $\gamma p \rightarrow \pi^+ \pi^- p$. Как упоминалось ранее, были разработаны и объединены две различные GAN. GAN моделирования детектора (DS-GAN) была обучена на PS-МС псевдоданных для изучения детекторных эффектов и затем интегрирована в архитектуру между генератором и дискриминатором Unfolded GAN (UNF-GAN) для восстановления "истинных" данных на уровне вершин из реконструированных псевдоданных (REC).

4.3.1 Генеративно-состязательная сеть для моделирования детектора (DS-GAN)

Для учета эффектов детектора была разработана модель машинного обучения на основе условной генеративной состязательной сети (Conditional GAN [84]). Схематическое представление модели представлено на рис. 4.9. Данный подход включает обучение генератора GAN для моделирования размывающего эффекта детектора так, чтобы он генерировал синтетические REC-события на уровне детектора, используя входные данные в виде вектора случайного шума и GEN PS-МС событий. Принятые PS-МС GEN-события передаются через цепочку GEANT, чтобы получить REC-псевдоданные. Как предложено в [55], как

синтетические REC, так и REC-псевдоданные объединяются с исходными GEN-событиями и подаются на вход дискриминатору GAN для упрощения сходимости сети. В процессе обучения DS-GAN учится на уровне отдельных событий оценивать влияние детектора на исходные данные. По завершении успешного обучения генератор DS-GAN становится заменой реального моделирования детектора, которая будет интегрирована в архитектуру UNF-GAN.

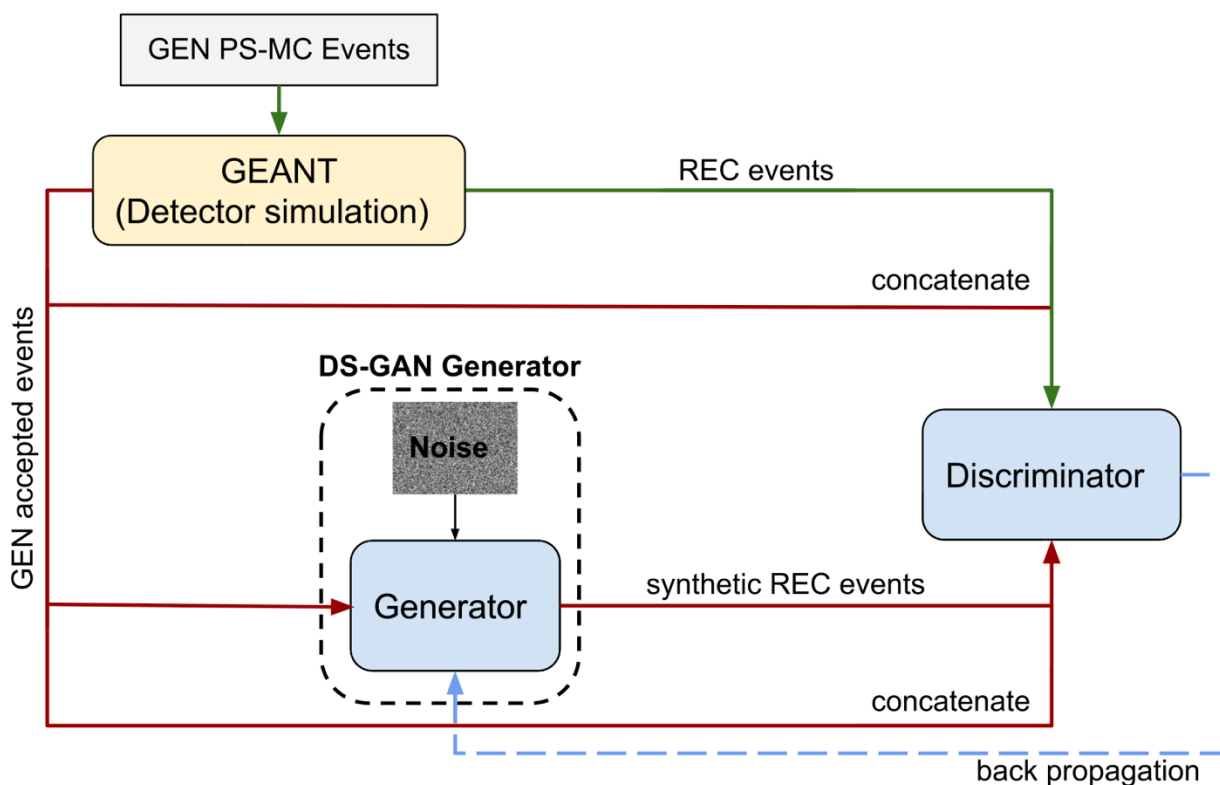


Рисунок 4.9 - Схематическое представление моделирования детектора с помощью машинного обучения на основе GAN (DS-GAN), где генератор GAN преобразует входные вершинные события GEN и шум в детекторные события REC. Обучение проводится на псевдоданных PS-MC, прошедших через моделирование GEANT. Синтетические REC-события и REC-псевдоданные объединяются с GEN PS-MC событиями и подаются дискриминатору.

Генератор DS-GAN, принимающий события GEN, получает на вход 100-мерный массив случайных значений со средним значением равным нулю и единичным стандартным отклонением. Генераторная сеть состоит из пяти скрытых слоев, каждый с 128 нейронами, с использованием функции активации Leaky ReLU. Финальный скрытый слой соединяется с выходным слоем из четырех нейронов, использующих линейную функцию для представления сгенерированных

признаков. После завершения обучения генератор DS-GAN учится преобразовывать принятые GEN-события в REC-события, эффективно моделируя размытие, вносимое детектором, согласно данным моделирования GEANT.

Дискриминатор представляет собой нейросеть с пятью скрытыми полносвязными слоями. Первые три слоя содержат по 256 нейронов, четвертый — 128 нейронов, пятый — 32 нейрона. Для всех слоев используется функция активации Leaky ReLU. Чтобы предотвратить переобучение в каждом скрытом слое, применяется 5%-ная вероятность дропаута (drop-out - метод, при котором случайным образом выключается часть нейронов) во время обучения. Последний скрытый слой полностью соединен с одним выходным нейроном, активированным линейной функцией, где "1" обозначает истинное событие, а "0" — сгенерированное. DS-GAN обучалась примерно на одном миллионе выборок событий фоторождения двух пионов в течение 8×10^5 эпох состязательного обучения. И генератор, и дискриминатор обучались с использованием оптимизатора ADAM с коэффициентом скорости обучения $\alpha = 10^{-5}$ и экспоненциальным затуханием градиента ($\beta^1 = 0.5, \beta^2 = 0.9$).

4.3.2 Генеративно-состязательная сеть для анфолдинга детекторных эффектов (UNF-GAN)

Процесс обучения UNF-GAN проиллюстрирован на рис. 4.10, на котором представлен вариант типовой структуры модели GAN, состоящей из условного генератора и дискриминатора. Генератор получает на вход энергию фотона, сгенерированную RE-MC, а также 100-мерный вектор белого шума со средним значением 0 и стандартным отклонением 1. Эта комбинация входных данных позволяет генератору, реализованному в виде глубокой нейронной сети, трансформировать шум и энергию фотона в минимальный набор переменных, эффективно описывающих процесс фоторождения двух пионов.

Чтобы достичь баланса между временем выполнения и сходимостью, сеть генератора состоит из семи скрытых слоев. Архитектура сети может быть представлена в виде вектора, где каждый элемент определяет число нейронов в соответствующем по порядку слое.

$$T = [16, 32, 64, 128, 256, 512, 1024, 4]. \quad (4.2)$$

Для всех слоев применяется функция активации ReLU. Последний скрытый слой полностью соединен с выходным слоем из четырех нейронов, использующих линейную активацию. Выходной слой представляет независимые переменные $M_{p\pi^-}^2, M_{\pi^+\pi^-}^2, t_{\pi^+}, \alpha_{[\pi^+p][\pi^-p']}$, которые специально выбраны для описания реакции.

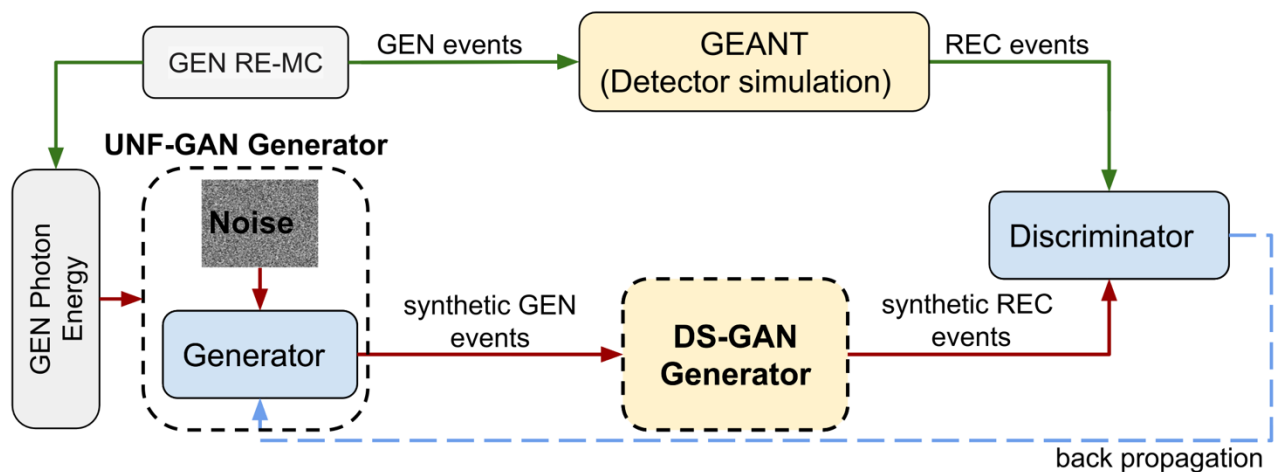


Рисунок 4.10 - Схематическое представление обучающей архитектуры UNF-GAN: UNF-GAN использует генератор, который преобразует энергию фотона на уровне GEN и случайный шум в синтетические характеристики событий GEN. Затем эти характеристики проходят через DS-GAN для включения детекторных эффектов и преобразуются в синтетические характеристики событий REC. Генератор обновляется с помощью градиентов, сформированных дискриминатором на основе глубокой нейронной сети, который сравнивает характеристики синтетических и эталонных REC-событий на детекторном уровне, полученных с помощью GEANT.

Сгенерированные GAN события на уровне вершины (GEN) затем передаются через DS-GAN для включения детекторных эффектов и сравниваются с REC-псевдоданными, полученными путем обработки RE-MC через GEANT. Обучение проводилось примерно на 4×10^5 событиях фоторождения двух пионов

в течение примерно 2×10^5 эпох состязательного обучения на каждую модель UNF-GAN. Параметры оптимизатора ADAM оставались неизменными и соответствовали тем, что использовались для DS-GAN. В процессе обучения генератор и дискриминатор ведут «состязательную борьбу», обновляя свои параметры. В результате генератор способен синтезировать REC-образцы, неотличимые от REC-псевдоданных. Это означает, что способность дискриминатора правильно классифицировать, является ли образец подлинным или синтетическим, приближается к случайному угадыванию.

4.3.3 Оценка неопределенности

С увеличением применения нейронных сетей в физическом анализе становится все более важным точно оценивать достоверность прогнозов моделей машинного обучения. Поскольку объем синтетических выборок может быть сколь угодно большим, статистическая неопределенность в этом случае пренебрежимо мала. Однако необходимо количественно оценить систематическую неопределенность, связанную с процессом обучения модели, и для этого была применена техника бутстрап-ресемплирования.

Для DS-GAN процедура включала независимое обучение 20 нейронных сетей. Каждая сеть обучалась на бутстрап-подвыборках исходного набора данных с повторными выборками, что привело к наборам данных одинакового размера, но с потенциально разными наблюдениями.

Для UNF-GAN применялся аналогичный процесс: 20 разных сетей обучались независимо, используя ту же технику бутстрап-ресемплирования. Более того, каждая из 20 UNF-GAN использовала разные версии DS-GAN из набора из 20 обученных моделей. Таким образом, оценка систематических неопределенностей для обеих сетей была унифицирована. Хотя использование большего количества бутстрап-подвыборок могло бы привести к более точным оценкам

неопределенности, было обнаружено, что обучение 20 моделей GAN дает достаточно стабильные результаты.

Важно отметить, что точное количество бутстрап-подвыборок может варьироваться в зависимости от особенностей задачи, объема доступных данных и требуемого уровня оценки неопределенности. В данном случае 20 бутстрап-выборки показали адекватность оценок неопределенностей, связанных с наблюдаемыми величинами. Кроме того, изменение архитектуры сети не требовалось, поскольку достигнутая сходимость, а также оцененные погрешности и неопределенности ясно указывают на то, что данная архитектура способна точно воспроизводить данные без введения дополнительных систематических ошибок.

4.4 Оценка качества моделирования процессов двухпионного фоторождения в эксклюзивных каналах

Далее приводится оценка производительности DS-GAN и UNF-GAN посредством сравнения синтетических данных с REC- и GEN-псевдоданными. Обозначения REC_{SYN} и GEN_{SYN} используются для синтетических данных на уровне детектора и вершинном уровне соответственно. Для визуального сравнения строились одномерные и двумерные гистограммы некоторых кинематических переменных. Чтобы показать, что корреляции правильно учтены, также строились распределения одной переменной при фиксированных значениях других переменных.

Поскольку синтетические данные генерируются с использованием бутстрап-процедуры, стандартное отклонение σ_{SYN} соответствует систематической неопределенности. Во всех наших результатах среднее значение μ_{SYN} показано сплошной линией вместе с областью ошибки шириной $\pm 1\sigma_{SYN}$, в то время как псевдоданные представлены точками с их статистической неопределенностью $\sigma_{pseudodata}$. Для количественной оценки степени соответствия между

синтетическими данными и псевдоданными были построены пул-распределения (pull distributions) для каждой ячейки, определяемые как:

$$pull = \frac{\mu_{SYN} - \mu_{pseudodata}}{\sqrt{\sigma_{SYN}^2 + \sigma_{pseudodata}^2}}, \quad (4.3)$$

где $\mu_{pseudodata}$ обозначает среднее по псевдоданным.

4.4.1 DS-GAN

DS-GAN обучается на четырех независимых переменных, включающих инвариантные массы $M_{p\pi^-}^2$ и $M_{\pi^+\pi^-}^2$, t_{π^+} и угол $\alpha_{[\pi^+p][\pi^-p']}$. Сравнение между распределениями PS-MC REC и REC_{SYN} показано на рис. 4.11. На рис. 4.12 представлено сравнение других значимых физических распределений, не использованных в обучении, являющихся функциями вышеупомянутых четырех переменных, а именно $M_{p\pi^+}^2$, t_{π^-} , t и $\cos\theta$. Количественное сравнение с использованием пул-распределений, представленных в нижней части каждого графика, подтверждает хорошее согласие данных в обоих случаях: большинство точек находится в пределах одного стандартного отклонения 1σ . Это указывает на то, что DS-GAN действительно способен обучаться эффектам детектора CLAS. Двумерные распределения из MC и синтетических данных представлены на рис. 4.13.

Абсолютное разрешение на импульс π^+ по псевдоданным (REC-GEN) представлено на рис. 4.14, а также его сравнение с синтетическими данными (REC_{SYN} - GEN). Два распределения демонстрируют хорошее согласие, что свидетельствует о корректном учете разрешения детектора. Аналогичные результаты получены для других кинематических переменных для всех частиц в конечном состоянии.

Данные сравнения демонстрируют способность DS-GAN изучать и воспроизводить эффекты детектора в многомерном пространстве, включая хвосты

распределений. Это подтверждает, что генеративные модели могут использоваться в качестве эффективных и быстрых заменителей более затратных по вычислениям GEANT-моделирований.

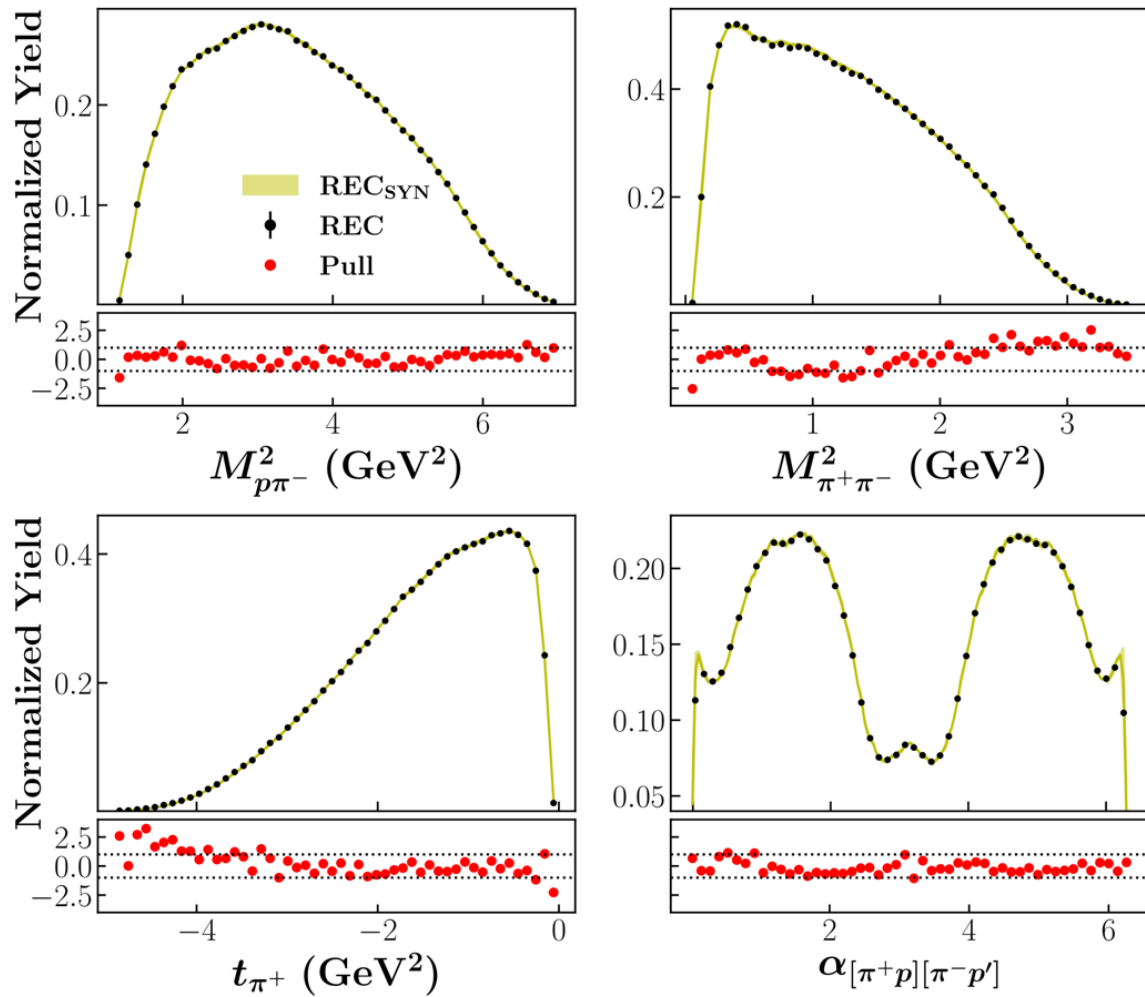


Рисунок 4.11 - Распределения обучающих переменных для REC-псевдоданных PS-MC (черные точки) сравниваются с синтетическими данными, сгенерированными DS-GAN (желтая полоса). Размер полосы отражает неопределенность, оцененную с помощью бутстрап-процедуры, пул-распределения (в нижней части) количественно характеризуют согласие между двумя наборами данных. Горизонтальные пунктирные линии соответствуют 1σ .

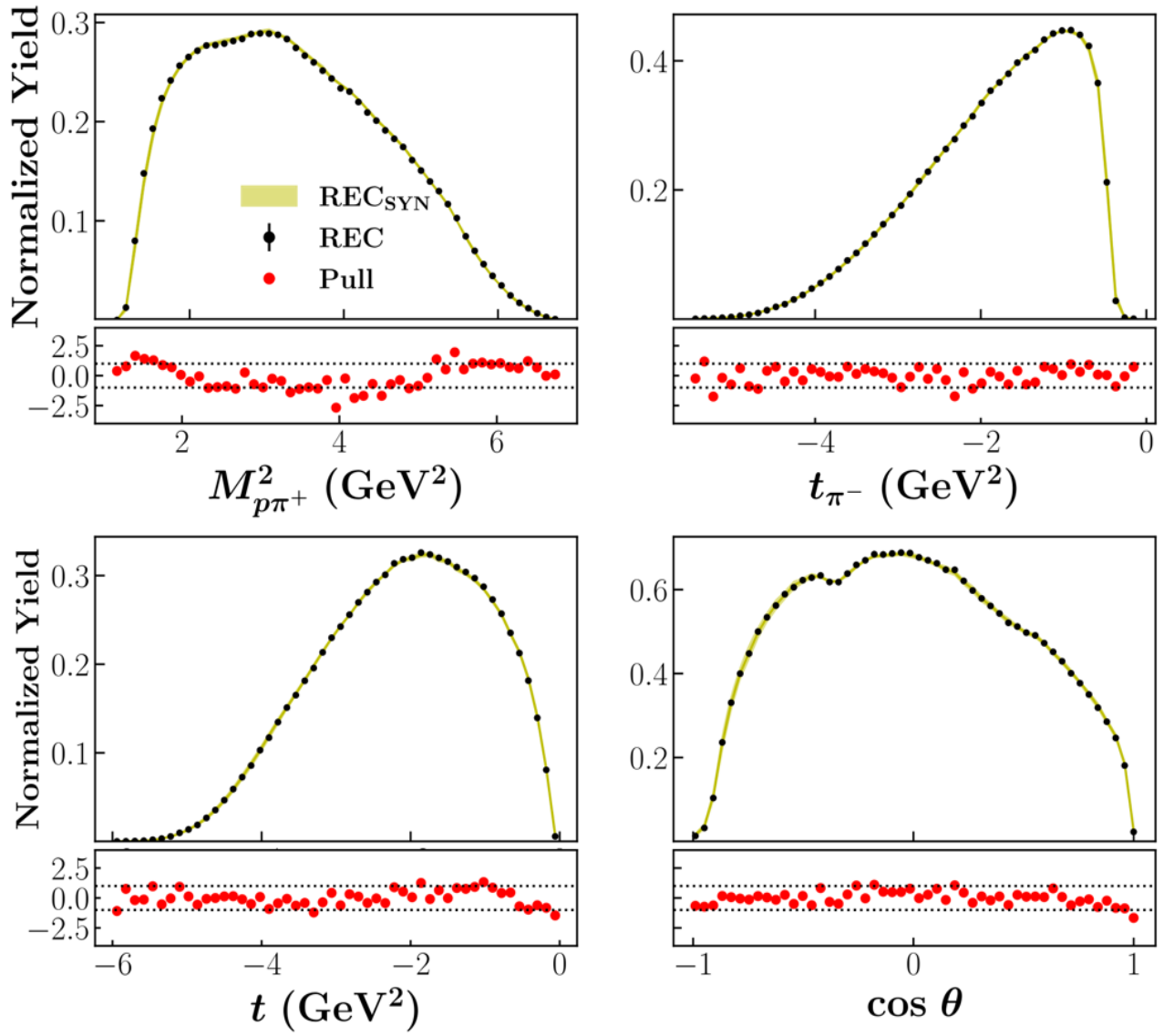


Рисунок 4.12 - Распределения переменных, выведенных и не участвовавших в обучении, для REC-псевдоданных PS-MC (черные точки) сравниваются с синтетическими данными, сгенерированными DS-GAN (желтая полоса). Размер полосы отражает неопределенность, оцененную с помощью бутстрап-процедуры, пул-распределения (в нижней части) количественно характеризуют согласие между двумя наборами данных. Горизонтальные пунктирные линии соответствуют 1σ .

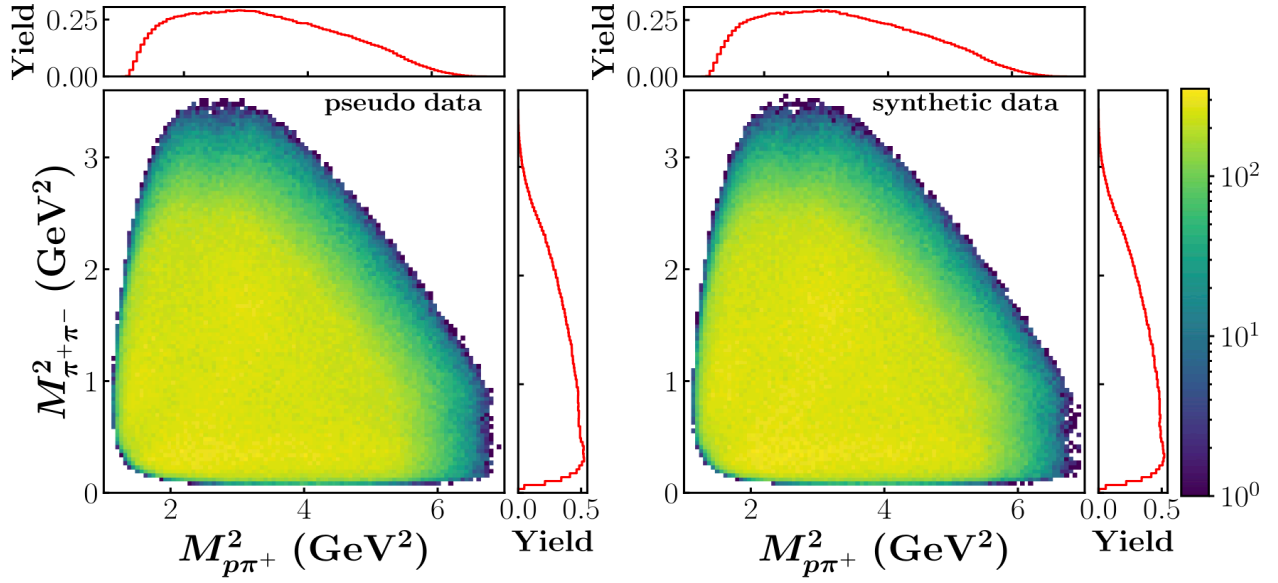


Рисунок 4.13 - 2D-распределения квадратов инвариантных масс и в REC-псевдоданных PS-MC (слева) и в синтетических данных, сгенерированных одной из моделей DS-GAN (справа).

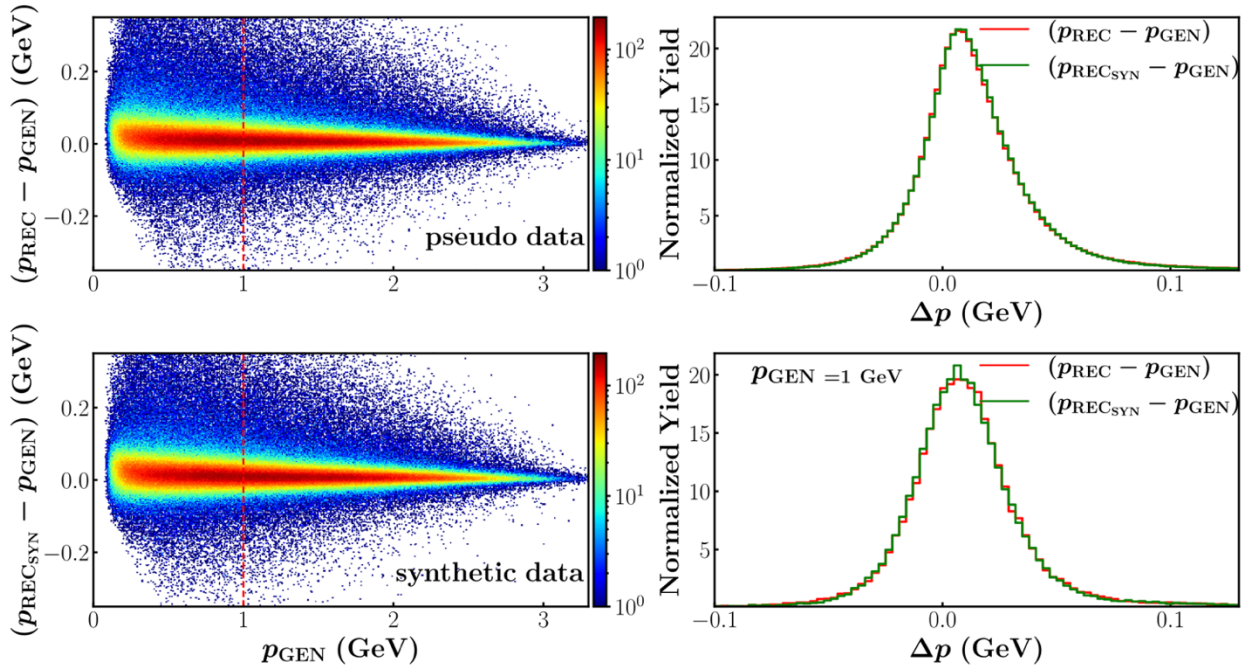


Рисунок 4.14 - Слева: сравнение абсолютного разрешения по импульсу $(p_{REC} - p_{GEN}$ и $p_{RECSYN} - p_{GEN}) \pi^+$ для REC-псевдоданных PS-MC (вверху) и синтетических данных DS-GAN (внизу) в зависимости от p_{GEN} . Справа: сравнение абсолютного разрешения по импульсу $(p_{REC} - p_{GEN}$ и $p_{RECSYN} - p_{GEN})$, интегрированного по p_{GEN} (вверху) или для узкого диапазона при $p_{GEN} = 1$ ГэВ (внизу).

4.4.2 UNF-GAN

Завершающий этап проверки заключается в использовании REC RE-МС-псевдоданных для обучения UNF-GAN, извлечении GEN_{SYN} распределений и их сравнении с GEN-псевдоданными. На рис. 4.15 продемонстрировано сравнение между GEN и GEN_{SYN} для четырех обучающих переменных. Можно видеть очень хорошее соответствие между псевдоданными и синтетическими данными на уровне вершины, несмотря на то, что UNF-GAN обучен на псевдоданных уровня детектора. Это однозначно демонстрирует успех процедуры анфолдинга.

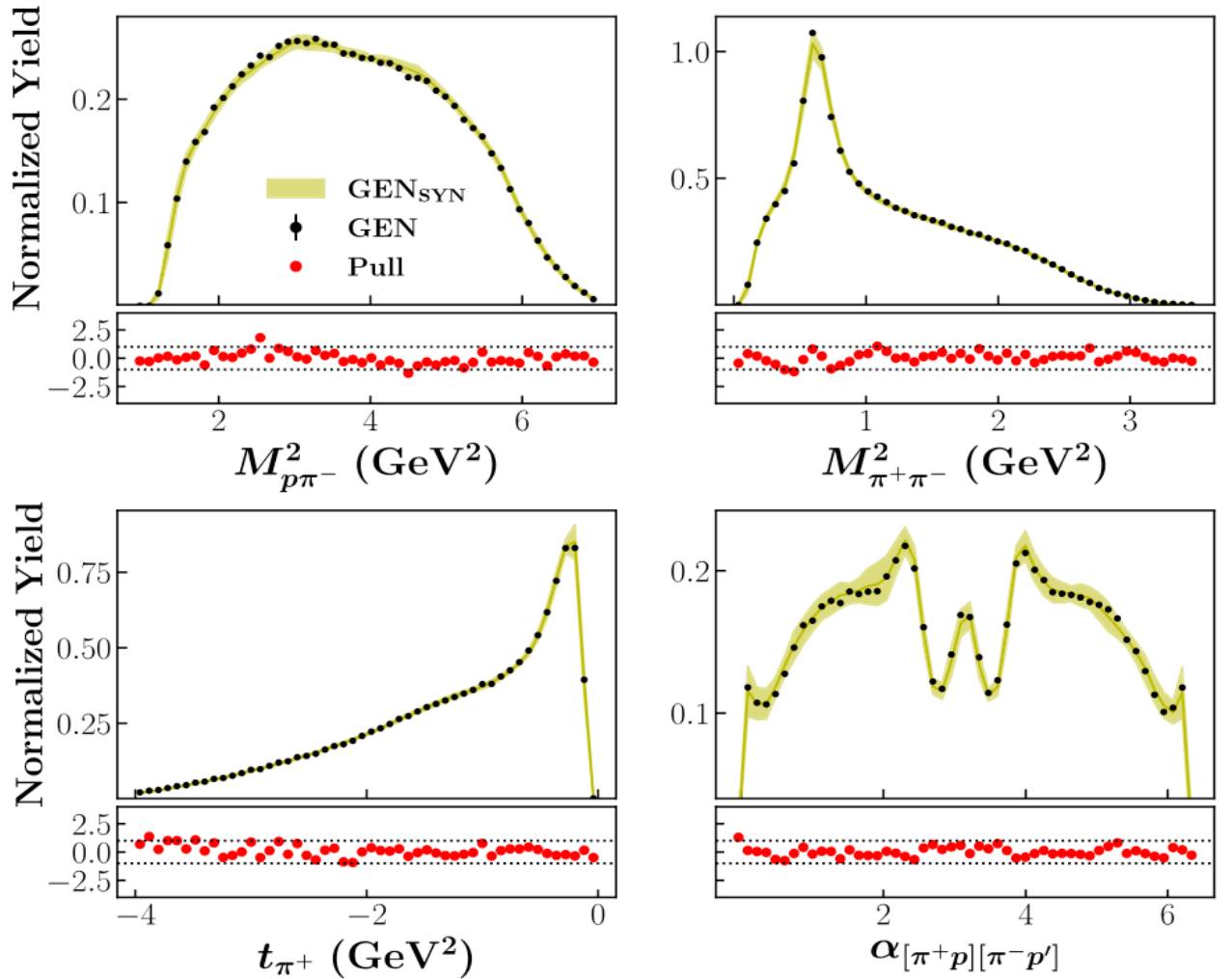


Рисунок 4.15 - Сравнение распределений GEN и GEN_{SYN} для четырех обучающих переменных, полученных с помощью UNF-GAN. Псевдоданные сгенерированы с использованием RE-МС.

Основная цель этого теста заключалась в том, чтобы доказать, что синтетические данные сохраняют корреляции исходных псевдоданных. Это действительно так, как видно из примеров двумерных распределений на рис. 4.16, содержащих сильные корреляции. Количественная оценка представлена в рассчитанных пул-распределениях на рис. 4.17. Самые высокие значения случайно распределены в двумерных распределениях (левый график), так что явных систематических тенденций нет. Это также подтверждается гистограммой пул-распределения (правый график), которая имеет нормальное распределение.

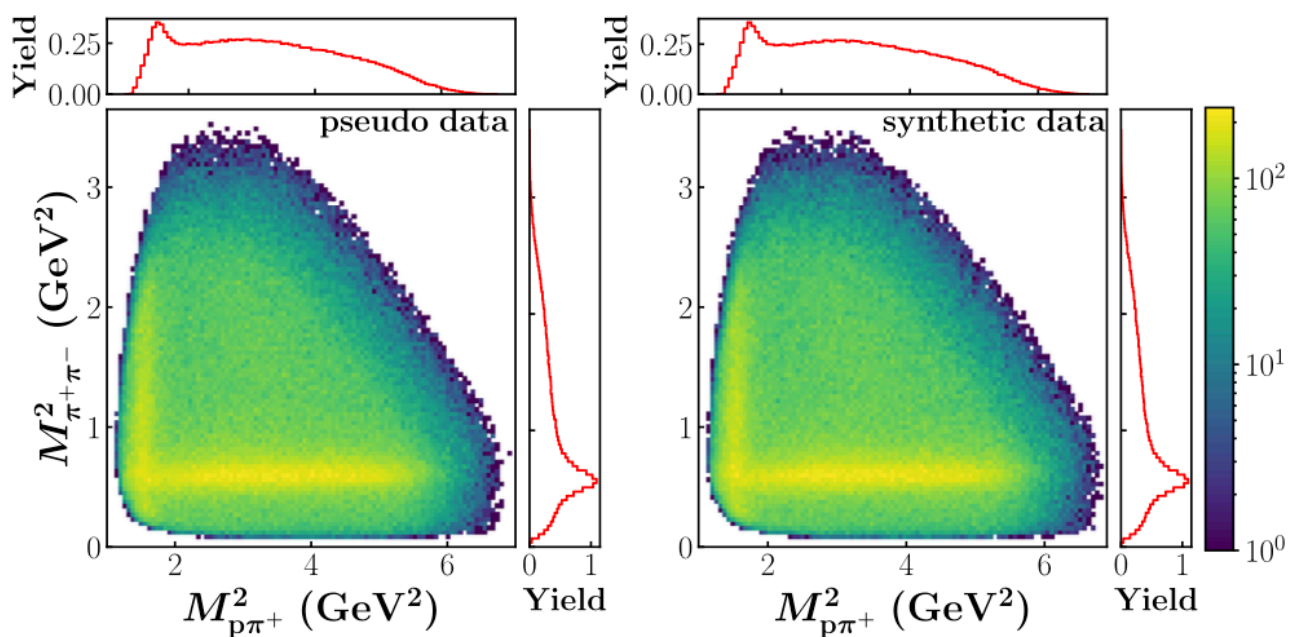


Рисунок 4.16 - 2D-распределения квадратов инвариантных масс $\pi^+\pi^-$ и $p\pi^+$ в GEN-псевдоданных RE-MC (слева) и в синтетических данных, сгенерированных одним из 20 UNF-GAN (справа).

Хорошее согласие и сохранение корреляций остаются справедливыми и для производных кинематических переменных, не использованных в обучении. Примеры представлены на рис. 4.18 для инвариантных переменных. Важно отметить, что в лабораторной системе GEN псевдоданные демонстрируют резкие особенности из-за ограничений детектора. Эти особенности не могут быть правильно воспроизведены GAN, который обучается на инвариантных переменных. Тем не менее, это приводит к локальным расхождениям менее 2σ в

одномерных проекциях. При необходимости в более точном соответствии переменные лабораторной системы могут быть добавлены в набор данных для обучения. На рис. 4.19 представлены одномерные гистограммы для фиксированного среза других переменных. Успех этого теста показывает, что корреляции, лежащие в основе дифференциального сечения рассеяния, правильно воспроизводятся в синтетических наборах данных.

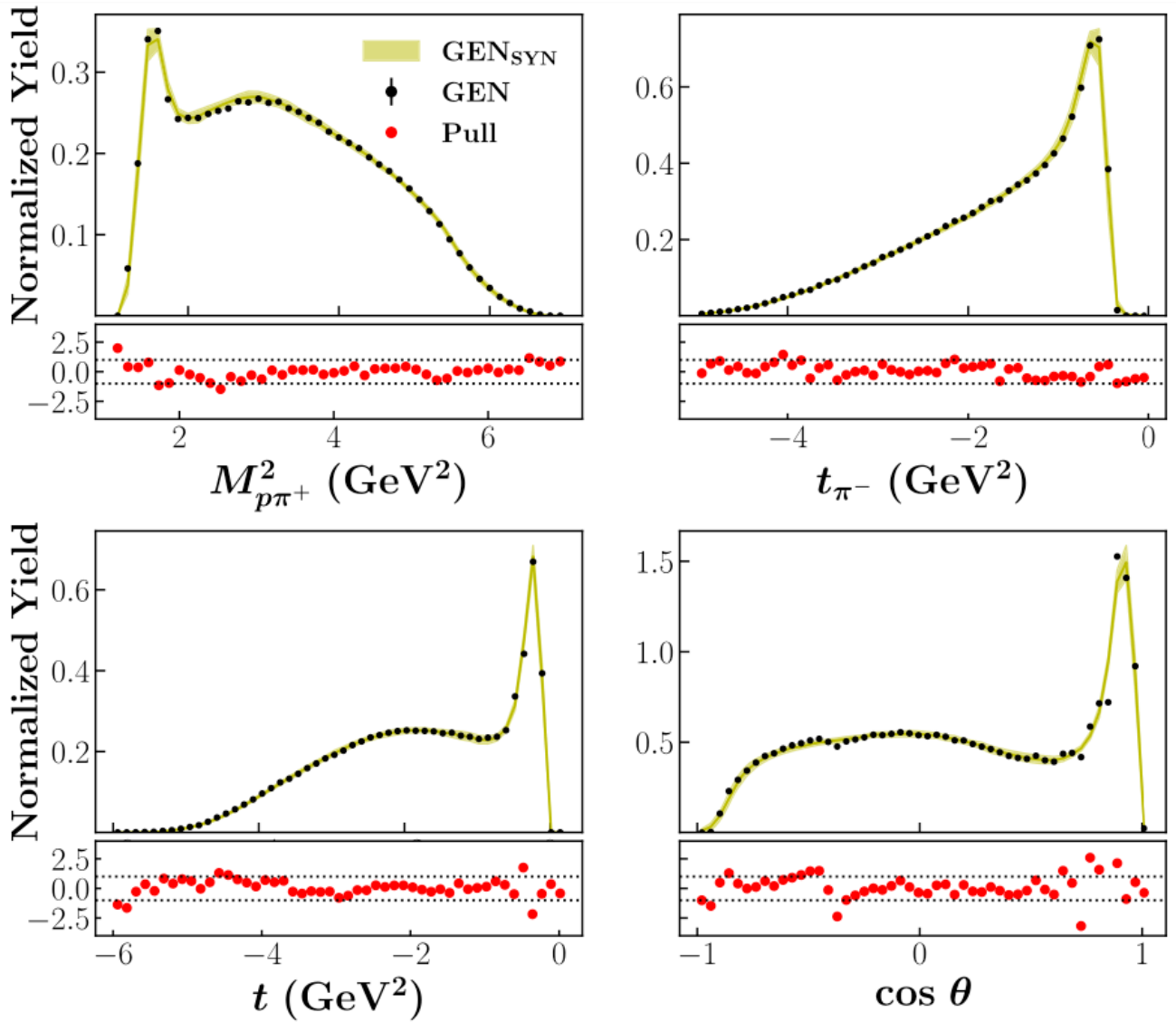


Рисунок 4.17 - Сравнение распределений GEN и GEN_{SYN} для четырех производных переменных: $M_{p\pi^+}^2$, t_{π^-} , t , $\cos \theta$. События были сгенерированы с использованием RE-MC.

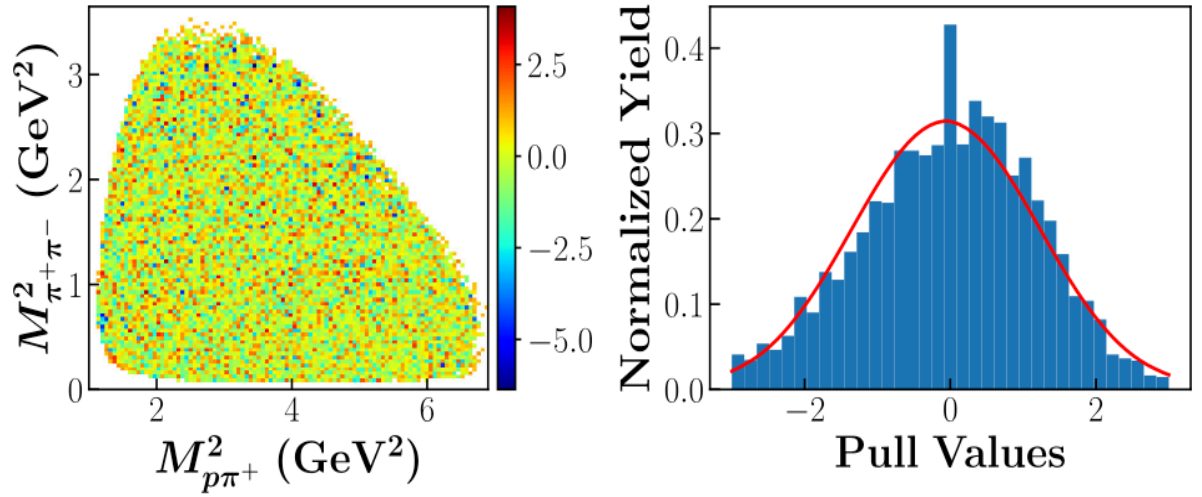


Рисунок 4.18 - Левый график: 2D пул-распределение для данных квадратов инвариантных масс $\pi^+\pi^-$ и $p\pi^+$. Цветовая палитра показывает значение пул. Правый график: гистограмма 2D пул-распределения с аппроксимацией гауссовым распределением, центрированным в нуле и имеющим стандартное отклонение, равное 1.3.

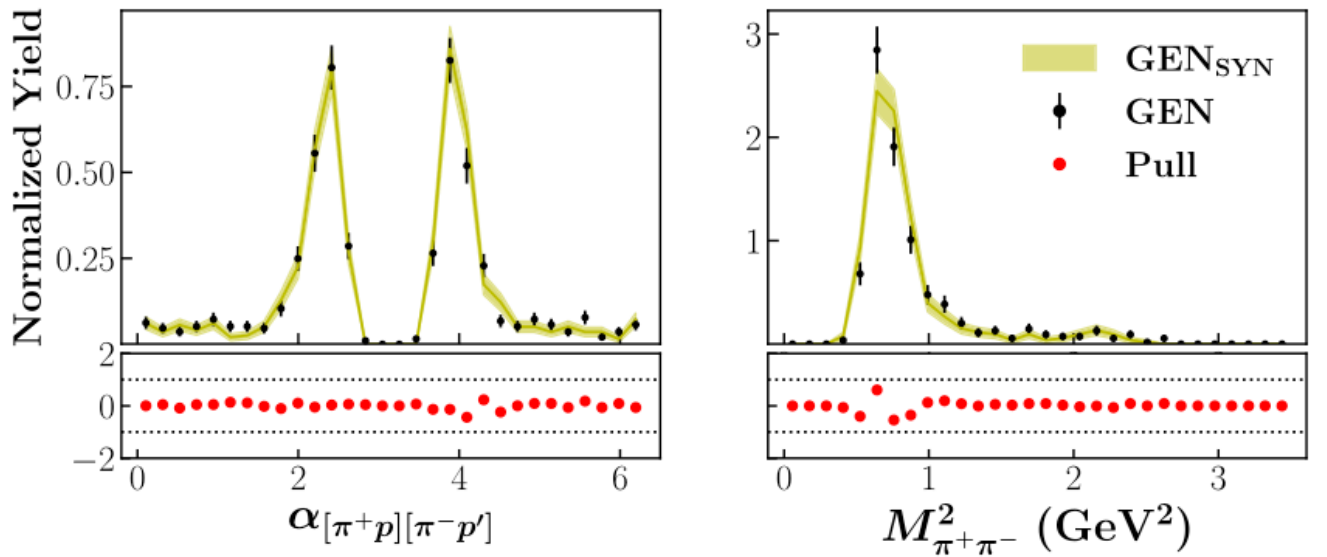


Рисунок 4.19 - 1D-гистограммы для фиксированного диапазона других переменных ($2.55 < W < 2.60$ ГэВ, $-0.7 < t_{\pi^+} < -0.3$ ГэВ², $2.5 < M_{p\pi^-}^2 < 3.3$ ГэВ²). Левый график: распределение $\alpha_{[\pi^+p][\pi^-p']}$ для $0.6 < M_{\pi^+\pi^-}^2 < 0.9$ ГэВ². Правый график: распределение $M_{\pi^+\pi^-}^2$ для $2.0 < \alpha_{[\pi^+p][\pi^-p']} < 2.5$.

Заключение

Основные результаты работы заключаются в следующем:

1. Разработан метод, основанный на полносвязных нейронных сетях, способный вычислять значения структурных функций и дифференциальных сечений рассеяния реакции $\gamma^* + p \rightarrow n + \pi^+$ в области инвариантной массы системы конечных адронов $1.1 < W < 2$ ГэВ, квадрата четырехимпульса виртуального фотона $0.3 < Q^2 < 4.1$ ГэВ² для различных значений энергии пучка электронов $E = 1.515, 5.499, 5.754$ ГэВ и телесного угла конечных адронов, близкого к 4π . Проведена процедура оценки погрешности вычислений, а также проведено сравнение с экспериментальными данными, показавшее, что результаты модели воспроизводят их со средней абсолютной ошибкой, равной 0.08 мкб/ср.
2. Получены значения структурных функций и дифференциальных сечений электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах в области квадрата четырехимпульса виртуального фотона $Q^2 < 5$ ГэВ² и инвариантной массы системы конечных адронов $W < 2.65$ ГэВ для различных значений энергии пучка электронов $E = 2.567, 4.056, 5.499$ ГэВ для телесного угла конечных адронов, близкого к 4π . Оценены погрешности полученных результатов. Проведено сравнение с экспериментальными данными, при котором установлено, что результаты модели согласуются с ними в пределах неопределенностей.
3. Разработан метод выявления многомерных корреляций в событиях реакции $\gamma + p \rightarrow p + \pi^+ + \pi^-$ в области инвариантной массы системы конечных адронов $W < 2.9$ ГэВ для телесного угла конечных адронов, близкого к 4π , основанный на генеративных алгоритмах машинного обучения. Независимая генеративно-состязательная нейронная сеть была создана и применена для учета детекторных эффектов. Оценки, полученные нейросетевым методом, демонстрируют согласие с данными, полученными при помощи Монте-Карло генератора в распределении pull-значений в пределах одного стандартного

отклонения. Продемонстрирована возможность восстановления четырехимпульсов частиц в реакции $\gamma + p \rightarrow p + \pi^+ + \pi^-$.

Полученные в данной работе результаты подтверждают способность методов машинного обучения улавливать сложные физические корреляции. Созданные алгоритмы способны улучшить наше понимание фото- и электровозбуждения резонансов.

Благодарности

В завершение хотелось бы выразить искреннюю благодарность всем, кто оказал поддержку в процессе работы над диссертацией.

Выражаю благодарность Анне Александровне Русовой за внимание, поддержку и участие, сыгравшие важную роль на различных этапах подготовки работы.

Особая признательность научному руководителю Исупову Евгению Леонидовичу за всестороннюю помощь, ценные советы на протяжении всей подготовки исследования и неизменную моральную поддержку.

Благодарю Виктора Ивановича Мокеева и участников коллаборации CLAS за профессиональные рекомендации и содействие, оказанные в ходе работы.

Выражаю глубокую благодарность Вячеславу Анатольевичу Ильину за ценные комментарии к работе, конструктивные замечания и полезные советы по ее оформлению и интерпретации полученных результатов.

Благодарю Лидию Ивановну Галанину за помощь в правильном оформлении работы и документов.

Также хочу поблагодарить сотрудников НИИ Ядерной Физики МГУ, в разной форме помогавших мне при подготовке диссертационной работы.

Список литературы

- [1] Gross F., Klempt E., Brodsky S.J., Buras A.J., Burkert V.D., Heinrich G., Jakobs K., Meyer C.A., Orginos K., Strickland M. et. al. 50 Years of quantum chromodynamics // European Physical Journal C. – 2023. – Vol. 83. – P. 1125. DOI: 10.1140/epjc/s10052-023-11949-2.
- [2] Carman D.S., Gothe R.W., Mokeev V.I., Roberts C.D. Nucleon Resonance Electroexcitation Amplitudes and Emergent Hadron Mass // Particles. – 2023. – Vol. 6. – № 1. – P. 416–439. DOI: 10.3390/particles6010023.
- [3] Скородумина Ю.А., Буркерт В.Д., Головач Е.Н., Готе Р.В., Исупов Е.Л., Ишханов Б.С., Мокеев В.И., Федотов Г.В. Нуклонные резонансы в эксклюзивных реакциях фото- и электророждения мезонов // Вестник Московского университета. Серия 3: Физика. Астрономия. – 2015. – № 6. DOI: 10.3103/S002713491506017X.
- [4] Crede V., Roberts W. Progress towards understanding baryon resonances // Reports on Progress in Physics. – 2013. – Vol. 76. – № 7. – P. 076301. DOI: 10.1088/0034-4885/76/7/076301.
- [5] Aznauryan I.G., Burkert V.D. Electroexcitation of nucleon resonances // Progress in Particle and Nuclear Physics. – 2012. – Vol. 67. – № 1. – P. 1–54. DOI: 10.1016/j.pnpnp.2011.08.001.
- [6] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain // Psychological Review. – 1958. – Vol. 65. – № 6. – P. 386–408. DOI: 10.1037/h0042519.

- [7] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. Attention Is All You Need // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 5998–6008.
- [8] Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities // Proceedings of the National Academy of Sciences of the United States of America. – 1982. – Vol. 79. – № 8. – P. 2554–2558. DOI: 10.1073/pnas.79.8.2554.
- [9] Ackley D.H., Hinton G.E., Sejnowski T.J. A learning algorithm for Boltzmann machines // Cognitive Science. – 1985. – Vol. 9. – № 1. – P. 147–169. DOI: 10.1207/s15516709cog0901_7.
- [10] Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A. et. al. Highly accurate protein structure prediction with AlphaFold // Nature. – 2021. – Vol. 596. – № 7873. – P. 583–589. DOI: 10.1038/s41586-021-03819-2.
- [11] Alghamdi T., Alanazi Y., Battaglieri M., Bibrzycki L., Golda A.V., Hiller Blin A. N., Isupov E.L., Li Y., Marsicano L., Melnitchouk W., Mokeev V.I., Montana G., Pilloni A., Sato N., Szczepaniak A. P., Vittorini T. Toward a generative modelling analysis of CLAS exclusive 2π photoproduction // Physical Review D. – 2023. – Vol. 108, № 9. – P. 094030. DOI: 10.1103/PhysRevD.108.094030.
- [12] Thomadakis P., Angelopoulos A., Gavalian G., Chrisochoides N. Using machine learning for particle track identification in the CLAS12 detector // Computer Physics Communications. – 2022. – Vol. 276. – P. 108360. DOI: 10.1016/j.cpc.2022.108360.
- [13] Официальный сайт Национальной ускорительной лаборатории имени Томаса Джефферсона – URL: <https://www.jlab.org/>

- [14] CLAS Physics Database – URL: <https://clas.sinp.msu.ru/cgi-bin/jlab/db.cgi>
- [15] Cutkosky R.E., Forsyth C.P., Hendrick R.E., Kelly R.L. Pion-nucleon partial-wave amplitudes // *Physical Review D*. – 1979. – Vol. 20. – № 11. – P. 2839–2853. DOI: 10.1103/PhysRevD.20.2839.
- [16] Amaldi E., Fubini S., Furlan G. Pion-Electroproduction: Electroproduction at Low Energy and Hadron Form Factors // *Springer Tracts in Modern Physics*. – Berlin; Heidelberg: Springer, – 1979. – Vol. 83. – P. 1–162. DOI: 10.1007/BFb0048208.
- [17] Голда А.В., Исупов Е.Л., Русова А.А., Рыжков В.Б. Исследование спектра и структуры нуклонных резонансов в коллаборации МГУ/JLAB // *Вестник Московского университета. Серия 3: Физика, астрономия* – 2026. – № 81. DOI: 10.55959/MSU0579-9392.81.2610204.
- [18] Knöchlein G., Drechsel D., Tiator L. Photo- and electroproduction of eta mesons // *Zeitschrift für Physik A Hadrons and Nuclei*. – 1995. – Vol. 352. – № 3. – P. 327–343. DOI: 10.1007/BF01289506.
- [19] Aznauryan I.G., Burkert V.D. Electroexcitation of nucleon resonances // *Progress in Particle and Nuclear Physics* – 2012. – Vol. 67. – № 1. – P. 1-54. DOI: 10.1016/j.pnpnp.2011.08.001.
- [20] Burkert V.D., Mokeev V.I., Ishkhanov B.S. The Nucleon Resonance Structure from the $\pi^+\pi^-p$ Electroproduction Reaction off Protons // *Moscow University Physics Bulletin*. – 2019. – Vol. 74. – № 3. – P. 243–255. DOI: 10.3103/S0027134919030032.
- [21] Borisov V., Leemann T., Seßler K., Haug J., Pawelczyk M., Kasneci G. Deep Neural Networks and Tabular Data: A Survey // *IEEE Transactions on Neural Networks and Learning Systems*. – 2024. – Vol. 35. – № 6. – P. 7499–7519. DOI: 10.1109/TNNLS.2022.3229161.

- [22] Agarap A.F. Deep Learning using Rectified Linear Units (ReLU) // arXiv preprint. – 2018. – arXiv:1803.08375. DOI: 10.48550/arXiv.1803.08375.
- [23] Golda A.V., Rusova A.A., Isupov E.L. et al. Machine Learning Approach in the Prediction of Differential Cross Sections and Structure Functions of Single Pion Electroproduction in the Resonance Region. // Moscow University Physics Bulletin. – 2024. – Vol. 79, № 4. –P. S608-S621. DOI: 10.3103/S0027134924702060.
- [24] Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators // Neural Networks. – 1989. – Vol. 2. – № 5. – P. 359–366. DOI: 10.1016/0893-6080(89)90020-8.
- [25] Чистякова В.В., Голда А.В., Русова А.А., Исупов Е.Л. Машинное обучение в задаче предсказания дифференциальных сечений и структурных функций электророждения пионов в резонансной области // Ученые записки физического факультета Московского университета – 2025. – №2. URL: <https://uzmu.phys.msu.ru/abstract/2025/2/2520213/>.
- [26] PyTorch: Python DNN framework. – URL: <https://pytorch.org/>
- [27] Schmidhuber J. Deep Learning in Neural Networks: An Overview // Neural Networks. – 2015. – Vol. 61. – P. 85–117. DOI: 10.1016/j.neunet.2014.09.003.
- [28] Weights & Biases platform. – URL: <https://wandb.ai/>
- [29] Meng C., Seo S., Cao D., Griesemer S., Liu Y. When Physics Meets Machine Learning: A Survey of Physics-Informed Machine Learning // arXiv preprint. – 2022. – arXiv:2203.16797. DOI: 10.48550/arXiv.2203.16797.
- [30] Rueden L., Mayer S., Beckh K., Georgiev B., Giesselbach S., Heese R., Kirsch B., Pfrommer J., Pick A., Ramamurthy R. et al. Informed Machine Learning – A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems //

- IEEE Transactions on Knowledge and Data Engineering. – 2023. – Vol. 35. – № 1. – P. 614–633. DOI: 10.1109/TKDE.2021.3079836.
- [31] Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization // arXiv preprint. – 2014. – arXiv:1412.6980. DOI: 10.48550/arXiv.1412.6980.
- [32] Drechsel D., Kamalov S.S., Tiator L. Unitary Isobar Model – MAID2007 // European Physical Journal A. – 2007. – Vol. 34. – P. 69–97. DOI: 10.1140/epja/i2007-10490-6.
- [33] Голда А.В., Исупов Е.Л., Русова А.А., Чистякова В.В. Нейросетевое моделирование наблюдаемых реакций электророждения одиночного пиона на протоне в резонансной области // Вестник Московского Университета. – 2026. – №2. DOI: 10.55959/MSU0579-9392.81.2620204.
- [34] Hobart A., Niccolai S., Čuić M., Stepanyan S., Biselli A.S., Burkert V.D., Carman D.S., Chatagnon P., Dutrieux H., Fegan S. et al. First Measurement of Deeply Virtual Compton Scattering on the Neutron with Detection of the Active Neutron // Physical Review Letters. – 2024. – Vol. 133. – № 21. – P. 211903. DOI: 10.1103/PhysRevLett.133.211903.
- [35] Ball R.D., Del Debbio L., Forte S., Guffanti A., Latorre J.I., Piccione A., Rojo J., Ubiali M. A determination of parton distributions with faithful uncertainty estimation // Nuclear Physics B. – 2009. – Vol. 809. – № 1–2. – P. 1–63. DOI: 10.1016/j.nuclphysb.2008.09.037.
- [36] Ball R.D., Del Debbio L., Forte S., Guffanti A., Latorre J.I., Rojo J., Ubiali M. A first unbiased global NLO determination of parton distributions and their uncertainties // Nuclear Physics B. – 2010. – Vol. 838. – № 1–2. – P. 136–206. DOI: 10.1016/j.nuclphysb.2010.05.008.

- [37] Голда А.В., Голубенко А.А., Давыдов М.М., Исупов Е.Л., Мокеев В.И., Савкин С.А., Чистякова В.В. Оценка сечений эксклюзивных каналов электророждения $K^+\Lambda$ и $K^+\Sigma^0$ на протонах из данных детектора CLAS. // Вестник Московского университета. Серия 3: Физика. Астрономия – 2024. – № 4, – С. 2440201. DOI: 10.55959/MSU0579-9392.79.2440201.
- [38] Mokeev V.I., Carman D.S. Photo- and Electrocouplings of Nucleon Resonances // Few-Body Systems. – 2022. – Vol. 63. – № 3. – P. 59. DOI: 10.1007/s00601-022-01760-2.
- [39] Brodsky S.J., Burkert V.D., Carman D.S., Chen J.P., Detmold W., Dudek J.J., Gothe R.W., Hatta Y., Hyde C.E., Meziani Z.-E. et al. Strong QCD from Hadron Structure Experiments // International Journal of Modern Physics E. – 2020. – Vol. 29. – № 8. – P. 2030006. DOI: 10.1142/S0218301320300064.
- [40] Carman D.S., Joo K., Mokeev V.I. Strong QCD Insights from Excited Nucleon Structure Studies with CLAS and CLAS12 // Few-Body Systems. – 2020. – Vol. 61. – № 3. – P. 29. DOI: 10.1007/s00601-020-01563-3.
- [41] Mokeev V.I., Aznauryan I.G., Burkert V.D., Gothe R.W. Recent results on the nucleon resonance spectrum and structure from the CLAS detector // EPJ Web of Conferences. – 2016. – Vol. 113. – P. 01013. DOI: 10.1051/epjconf/201611301013.
- [42] Ambrozewicz P., Carman D.S., Feuerbach R.J., Mestayer M.D., Raue B.A., Schumacher R.A., Tkabladze A., Adams G., Adhikari K.P., Amarian M.J. et al. Separated structure functions for the exclusive electroproduction of $K^+\Lambda$ and $K^+\Sigma^0$ final states // Physical Review C. – 2007. – Vol. 75. – № 4. – P. 045203. DOI: 10.1103/PhysRevC.75.045203.
- [43] Carman D.S., Park K., Raue B.A., Adhikari K.P., Adikaram D., Aghasyan M., Amarian M.J., Anderson M.D., Anghinolfi M., Avakian H. et al. Separated structure functions for exclusive $K^+\Lambda$ and $K^+\Sigma^0$ electroproduction at 5.5 GeV measured with

- CLAS // Physical Review C. – 2013. – Vol. 87. – № 2. – P. 025204. DOI: 10.1103/PhysRevC.87.025204.
- [44] Miller G., Bloom E.D., Buschhorn G., Coward D.H., DeStaebler H., Drees J., Mo L.W., Taylor R.E. Inelastic electron-proton scattering at large momentum transfers and the inelastic structure functions of the proton // Physical Review D. – 1972. – Vol. 5. – № 3. – P. 528–544. DOI: 10.1103/PhysRevD.5.528.
- [45] Lanza L., D’Angelo A. KY electroproduction at CLAS12 // Il Nuovo Cimento C. – 2021. – Vol. 44. – № 2–3. – P. 51. DOI: 10.1393/ncc/i2021-21051-4.
- [46] Mokeev V.I., Burkert V.D., Carman D.S., Elouadrhiri L., Golovatch E., Gothe R.W., Hicks K., Ishkhanov B.S., Isupov E.L., Joo K. et al. Evidence for the $N'(1720)3/2^+$ nucleon resonance from combined studies of CLAS $\pi^+\pi^-p$ photo- and electroproduction data // Physics Letters B. – 2020. – Vol. 805. – P. 135457. DOI: 10.1016/j.physletb.2020.135457.
- [47] Melnitchouk W., Ent R., Keppel C.E. Quark–hadron duality in electron scattering // Physics Reports. – 2005. – Vol. 406. – № 3–4. – P. 127–301. DOI: 10.1016/j.physrep.2004.10.004.
- [48] Репозиторий исходного кода для экстраполяции и интерполяции дифференциальных сечений рассеяния. – URL: https://github.com/maksaska/Diff_cros_strange/tree/python-version.
- [49] Веб-сайт с интерфейсом для расчета дифференциальных сечений рассеяния. – URL: <https://clas.sinp.msu.ru/~maksaska/>.
- [50] Battaglieri M., De Vita R., Kubarovsky V., Price J., Stoler P., Weygand D. Study of pentaquark states in photoproduction off protons: Jefferson Lab experiment proposal PR04-017. – 2004. – URL: http://www.jlab.org/exp_prog/proposals/04/PR04-017.pdf

- [51] Andreassen A., Komiske P.T., Metodiev E.M., Nachman B., Thaler J. OmniFold: A Method to Simultaneously Unfold All Observables // Physical Review Letters. – 2020. – Vol. 124. – № 18. – P. 182001. DOI: 10.1103/PhysRevLett.124.182001.
- [52] Andreev V., Arratia M., Baghdasaryan A., Baty A., Begzsuren K., Belousov A., Bolz A., Boudry V., Brandt G., Britzger D. et al. Measurement of Lepton-Jet Correlation in Deep-Inelastic Scattering with the H1 Detector Using Machine Learning for Unfolding // Physical Review Letters. – 2022. – Vol. 128. – № 13. – P. 132002. DOI: 10.1103/PhysRevLett.128.132002.
- [53] Andreev V., Arratia M., Baghdasaryan A., Baty A., Begzsuren K., Bolz A., Boudry V., Brandt G., Britzger D., Buniatyan A. et al. Unbinned Deep Learning Jet Substructure Measurement in High Q^2 ep Collisions at HERA // Physics Letters B. – 2023. – Vol. 844. – P. 138101. DOI: 10.1016/j.physletb.2023.138101.
- [54] Aaij R., Abdelmotteleb A.S.W., Abellán Beteta C., Abudinen F., Ackernley T., Adeva B., Adinolfi M., Afsharnia H., Aidala C.A., Aiola S. et al. Observation of New Resonances Decaying to $J/\psi K^+$ and $J/\psi \phi$ // Physical Review Letters. – 2021. – Vol. 127. – № 8. – P. 082001. DOI: 10.1103/PhysRevLett.127.082001.
- [55] Bellagente M., Butter A., Kasieczka G., Plehn T., Winterhalder R. How to GAN away Detector Effects // SciPost Physics. – 2020. – Vol. 8. – № 4. – P. 070. DOI: 10.21468/SciPostPhys.8.4.070.
- [56] Alanazi Y., Ambrozewicz P., Battaglieri M., Hiller Blin A.N., Kuchera M.P., Li Y., Liu T., McClellan R.E., Melnitchouk W., Pritchard E. et al. Machine learning-based event generator for electron-proton scattering // Physical Review D. – 2022. – Vol. 106. – № 9. – P. 096002. DOI: 10.1103/PhysRevD.106.096002.

- [57] Brun R., Bruyant F., Maire M., McPherson A.C., Zancarini P. GEANT 3: User's Guide Geant 3.10, Geant 3.11; rev. version. – Geneva: CERN, 1987. – URL: <https://cds.cern.ch/record/1119728>
- [58] Capstick S., Roberts W. Quark Models of Baryon Masses and Decays // Progress in Particle and Nuclear Physics. – 2000. – Vol. 45. – P. S241–S331. arXiv/0008028. DOI: 10.48550/arXiv.nucl-th/0008028.
- [59] Klempt E., Richard J.-M. Baryon spectroscopy // Reviews of Modern Physics. – 2010. – Vol. 82. – № 2. – P. 1095–1153. DOI: 10.1103/RevModPhys.82.1095.
- [60] Giannini M.M., Santopinto E. The Hypercentral Constituent Quark Model and Its Application to Baryon Properties // Chinese Journal of Physics. – 2015. – Vol. 53. – № 1. – P. 020301-1–020301-75. arXiv:1501.03722.
- [61] Qin S.-X., Roberts C.D., Schmidt S.M. Spectrum of light- and heavy-baryons // Few-Body Systems. – 2019. – Vol. 60. – № 2. – P. 26. DOI: 10.1007/s00601-019-1488-x.
- [62] Chen C., Krein G., Roberts C.D., Schmidt S.M., Segovia J. Spectrum and structure of octet and decuplet baryons and their positive-parity excitations // Physical Review D. – 2019. – Vol. 100. – № 5. – P. 054009. DOI: 10.1103/PhysRevD.100.054009.
- [63] Edwards R.G., Dudek J.J., Richards D.G., Wallace S.J. Excited state baryon spectroscopy from lattice QCD // Physical Review D. – 2011. – Vol. 84. – № 7. – P. 074508. DOI: 10.1103/PhysRevD.84.074508.
- [64] Crede V., Roberts W. Progress towards understanding baryon resonances // Reports on Progress in Physics. – 2013. – Vol. 76. – № 7. – P. 076301. DOI: 10.1088/0034-4885/76/7/076301.

- [65] Ireland D.G., Pasyuk E., Strakovsky I.I. Photoproduction reactions and non-strange baryon spectroscopy // Progress in Particle and Nuclear Physics. – 2020. – Vol. 111. – P. 103752. DOI: 10.1016/j.pnpnp.2019.103752.
- [66] Beck R., Thoma U. Spectroscopy of baryon resonances // EPJ Web of Conferences. – 2017. – Vol. 134. – P. 02001. DOI: 10.1051/epjconf/201713402001.
- [67] Burkert V., Klempt E., Thoma U. Light-quark baryons // arXiv preprint. – 2022. – arXiv:2211.12906. DOI: 10.48550/arXiv.2211.12906.
- [68] Ripani M., Anghinolfi M., Battaglieri M., Bianchi N., Biselli A.S., Briscoe W.J., Brooks W.K., Burkert V.D., Calarco J.R., Carman D.S. et al. Measurement of $ep \rightarrow e'\pi^+\pi^-$ and baryon resonance analysis // Physical Review Letters. – 2003. – Vol. 91. – № 2. – P. 022002. DOI: 10.1103/PhysRevLett.91.022002.
- [69] Carver M., Celentano A., Hicks K., Marsicano L., Mathieu V., Pilloni A., Adhikari K.P., Gilfoyle G.P., Ireland D.G., Mokeev V.I. et al. Photoproduction of the $f_2(1270)$ meson using the CLAS detector // Physical Review Letters. – 2021. – Vol. 126. – № 8. – P. 082002. DOI: 10.1103/PhysRevLett.126.082002.
- [70] Mathieu V., Nys J., Fernández-Ramírez C., Hiller Blin A.N., Jackura A., Pilloni A., Szczepaniak A.P., Fox G. Structure of pion photoproduction amplitudes // Physical Review D. – 2018. – Vol. 98. – № 1. – P. 014041. DOI: 10.1103/PhysRevD.98.014041.
- [71] Mathieu V., Pilloni A., Albaladejo M., Bibrzycki L., Celentano A., Fernández-Ramírez C., Szczepaniak A.P. Exclusive tensor meson photoproduction // Physical Review D. – 2020. – Vol. 102. – № 1. – P. 014003. DOI: 10.1103/PhysRevD.102.014003.
- [72] Battaglieri M., De Vita R., Szczepaniak A.P., Adhikari K.P., Anghinolfi M., Amaryan M.J., Ball J., Bianchi N. et al. Photoproduction of $\pi^+\pi^-$ meson pairs on

- the proton // *Physical Review D*. – 2009. – Vol. 80. – № 7. – P. 072005. DOI: 10.1103/PhysRevD.80.072005.
- [73] Mestayer M.D., Carman D.S., Asavapibhop B., Barbosa F.J., Bosshard B., Branford D., Brooks W.K., Burkert V.D., Cords D.I., Crabb D.G. et al. The CLAS drift chamber system // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. – 2000. – Vol. 449. – № 1–2. – P. 81–111. DOI: 10.1016/S0168-9002(00)00151-0.
- [74] Sober D.I., Crannell H., Longhi A., Matthews S.K., O’Brien J.T., Berman B.L., Briscoe W.J., Cole P.L., Connelly J.P., Dodge W.R. et al. The bremsstrahlung tagged photon beam in Hall B at JLab // *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. – 2000. – Vol. 440. – № 2. – P. 263–284. DOI: 10.1016/S0168-9002(99)00784-6.
- [75] Whalley M.R. The Durham-RAL high energy physics database — HEPDATA // *Computer Physics Communications*. – 1989. – Vol. 57. – № 1–3. – P. 536–537. DOI: 10.1016/0010-4655(89)90282-8.
- [76] Gsim user’s guide. – URL: <https://www.jlab.org/Hall-B/document/gsim/userguide.html>
- [77] Freyberger A., Ito M. CLAS offline code management: Technical Report CLAS-NOTE 97-0003. – CLAS Collaboration. – URL: ftp://ftp.cebaf.gov/pub/clas/notes/clas_notes97.html
- [78] Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative Adversarial Nets // *Advances in Neural Information Processing Systems*. – 2014. – Vol. 27. – arXiv:1406.2661. DOI: 10.48550/arXiv.1406.2661.

- [79] Dash A., Ye J., Wang G. A Review of Generative Adversarial Networks (GANs) and Its Applications in a Wide Variety of Disciplines: From Medical to Remote Sensing // IEEE Access. – 2024. – Vol. 12. – P. 18330–18357. DOI: 10.1109/ACCESS.2023.3346273.
- [80] Bang D., Shim H. MGGAN: Solving Mode Collapse using Manifold Guided Training // arXiv preprint. – 2018. – arXiv:1804.04391. DOI: 10.48550/arXiv.1804.04391.
- [81] Gao C., Höche S., Isaacson J., Krause C., Schulz H. Event generation with normalizing flows // Physical Review D. – 2020. – Vol. 101. – № 7. – P. 076002. DOI: 10.1103/PhysRevD.101.076002.
- [82] Musella P., Pandolfi F. Fast and Accurate Simulation of Particle Detectors Using Generative Adversarial Networks // Computing and Software for Big Science. – 2018. – Vol. 2. – № 1. – P. 8. DOI: 10.1007/s41781-018-0015-y.
- [83] Alanazi Y., Sato N., Liu T., Melnitchouk W., Ambrozewicz P., Hauenstein F., Kuchera M.P., Pritchard E., Robertson M., Strauss R. et al. Simulation of Electron-Proton Scattering Events by a Feature-Augmented and Transformed Generative Adversarial Network (FAT-GAN) // Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. – 2021. – P. 2126–2133. DOI: 10.24963/ijcai.2021/293.
- [84] Mirza M., Osindero S. Conditional Generative Adversarial Nets // arXiv preprint. – 2014. – arXiv:1411.1784. DOI: 10.48550/arXiv.1411.1784.