

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ «МИФИ»
(НИЯУ МИФИ)

ИНСТИТУТ ЯДЕРНОЙ ФИЗИКИ И ТЕХНОЛОГИЙ
КАФЕДРА №40 «ФИЗИКА ЭЛЕМЕНТАРНЫХ ЧАСТИЦ»

УДК 531.3, 539.1.05

На правах рукописи

КОСИЦЫН АРСЕНИЙ МАКСИМОВИЧ

**ИЗУЧЕНИЕ РАСПАДОВ БОЗОНА ХИГГСА НА
НЕРЕГИСТРИРУЕМЫЕ ЧАСТИЦЫ В ЭКСПЕРИМЕНТЕ СЕРС**

Направления подготовки:

09.04.04 «Программная инженерия»,

14.04.02 «Ядерная физика и технологии»

Диссертация на соискание степени магистра

Научный руководитель,
к.ф.-м.н.

_____ Е. Ю. Солдатов

Научный консультант,
к.т.н.

_____ М. В. Чадеева

Москва 2026

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА МАГИСТРА

ИЗУЧЕНИЕ РАСПАДОВ БОЗОНА ХИГГСА НА НЕРЕГИСТРИРУЕМЫЕ ЧАСТИЦЫ В ЭКСПЕРИМЕНТЕ СЕРС

Студент	_____ А. М. Косицын
Научный руководитель, к.ф.-м.н.	_____ Е. Ю. Солдатов
Научный консультант, к.т.н.	_____ М. В. Чадеева
Рецензент 1, к.ф.-м.н.	_____ М. С. Лицарев
Рецензент 2, к.ф.-м.н.	_____ А. Г. Мягков
Секретарь ГЭК, к.ф.-м.н.	_____ Е. Ю. Солдатов
Зав. каф. №40, д.ф.-м.н., проф.	_____ М. Д. Скорохватов
Рук. учеб. прог., к.ф.-м.н.	_____ Е. Ю. Солдатов

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 Обзор эксперимента СЕРС	7
1.1 Физическая программа и сценарии работы СЕРС	8
1.2 Базовый детектор СЕРС	10
2 Программное обеспечение	15
2.1 Архитектура и базовые компоненты	15
2.2 Идентификация изолированных частиц	16
2.3 Пакет FastJet для анализа струй	19
3 Анализ	21
3.1 Моделирование	21
3.2 Предварительные отборы	23
3.3 Основные отборы	24
3.4 Сравнение файлов с чистым и инклюзивным сигналом	27
3.5 Результаты аппроксимации	29
3.6 BDT-классификатор для подавления фоновых процессов	30
ЗАКЛЮЧЕНИЕ	39
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	43

ВВЕДЕНИЕ

С момента открытия бозона Хиггса на Большом адронном коллайдере (LHC) его свойства интенсивно изучаются [1; 2]. На сегодняшний день измеренные характеристики бозона Хиггса хорошо согласуются с предсказаниями Стандартной Модели (СМ) (рисунок 1 и таблица 1). Тем не менее, СМ не содержит кандидата на роль тёмной материи, что может указывать на существование физики за пределами Стандартной Модели (BSM). Во многих таких моделях взаимодействие частиц тёмной материи с обычным веществом осуществляется исключительно через поле Хиггса [3]. В этом случае бозон Хиггса может распадаться на нерегистрируемые детектором частицы, что приводит к увеличению вероятности такого “невидимого” распада по сравнению с предсказанием СМ.

В рамках СМ невидимый распад бозона Хиггса возможен только через канал $H \rightarrow ZZ^* \rightarrow 4\nu$ (рис. 2), и его предсказанная относительная вероятность распада составляет всего $\sim 0.1\%$. Любое экспериментально наблюдаемое превышение этого значения стало бы прямым указанием на новую физику или рождение частиц тёмной материи. На LHC поиски невидимого распада уже проводились, и коллаборации ATLAS и CMS установили верхние пределы на $\text{BR}(H \rightarrow \text{invisible})$ на уровне 11% и 15% соответственно [4; 5]. Однако из-за высокого фона и сложной кинематики адронных столкновений достижение процентной точности на LHC, даже в режиме высокой светимости, остаётся крайне затруднительным.

В настоящей работе исследовался канал, в котором бозон Хиггса рождается в ассоциации с Z -бозоном, где Z -бозон распадается на кварк антикварковую пару, а бозон Хиггса на нерегистрируемые частицы. Относительная вероятность распада $Z \rightarrow q\bar{q}$ (около 70%) значительно больше, чем у процесса с распадом Z -бозона на заряженные лептоны. Однако данный канал имеет множество фоновых процессов, мимикрирующих под его топологию. Среди основных фонов можно выделить:

- двухфермионные процессы с двумя джетами в конечном состоянии ($e^+e^- \rightarrow q\bar{q}$);
- четырехфермионные процессы с двумя джетами в конечном

состоянии ($e^+e^- \rightarrow e\nu_e q\bar{q}$, $e^+e^- \rightarrow Ze^+e^- \rightarrow q\bar{q}e^+e^-$, $e^+e^- \rightarrow Z\nu\bar{\nu} \rightarrow q\bar{q}\nu\bar{\nu}$);

— четырехфермионные процессы с неправильно объединенными четырьмя джетами в конечном состоянии ($e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}q\bar{q}$, $e^+e^- \rightarrow ZZ \rightarrow q\bar{q}q\bar{q}$);

— процессы рождения бозона Хиггса при слиянии векторных бозонов ($e^+e^- \rightarrow \nu\bar{\nu}H(\text{inclusive})$).

Наиболее сложными фонами при реконструкции сигнального процесса являются $e^+e^- \rightarrow Z\nu\bar{\nu} \rightarrow q\bar{q}\nu\bar{\nu}$ и $e^+e^- \rightarrow \nu\bar{\nu}H(\text{inclusive})$.

Основной исследуемой переменной является масса отдачи. Используя реконструированные джеты от распада Z -бозона, можно восстановить массу бозона Хиггса методом массы отдачи [6]:

$$M_{\text{recoil}} = \sqrt{(\sqrt{s} - E_{j1j2})^2 - |\vec{p}_{j1j2}|^2}, \quad (1)$$

где E_{j1j2} и $\vec{p}_{j1j2}$ — полная энергия и трёхмерный импульс системы двух джетов, а $\sqrt{s} = 240$ ГэВ — энергия в системе центра масс. Данный метод позволяет реконструировать массу невидимой системы, не требуя детектирования продуктов распада бозона Хиггса, что критически важно для поиска канала $H \rightarrow \text{invisible}$.

Таблица 1 – Относительные вероятности и относительные погрешности для распада бозона Хиггса с массой $m_H = 125$ ГэВ в Стандартной Модели [7]

Канал распада	Вероятность распада	Относительная погрешность
$H \rightarrow \gamma\gamma$	2.27×10^{-3}	2.1%
$H \rightarrow ZZ$	2.62×10^{-2}	1.5%
$H \rightarrow W^+W^-$	2.14×10^{-1}	1.5%
$H \rightarrow \tau^+\tau^-$	6.27×10^{-2}	1.6%
$H \rightarrow b\bar{b}$	5.82×10^{-1}	+1.2% / – 1.3%
$H \rightarrow c\bar{c}$	2.89×10^{-2}	5.8%
$H \rightarrow Z\gamma$	1.53×10^{-3}	5.5%
$H \rightarrow \mu^+\mu^-$	2.18×10^{-4}	1.7%

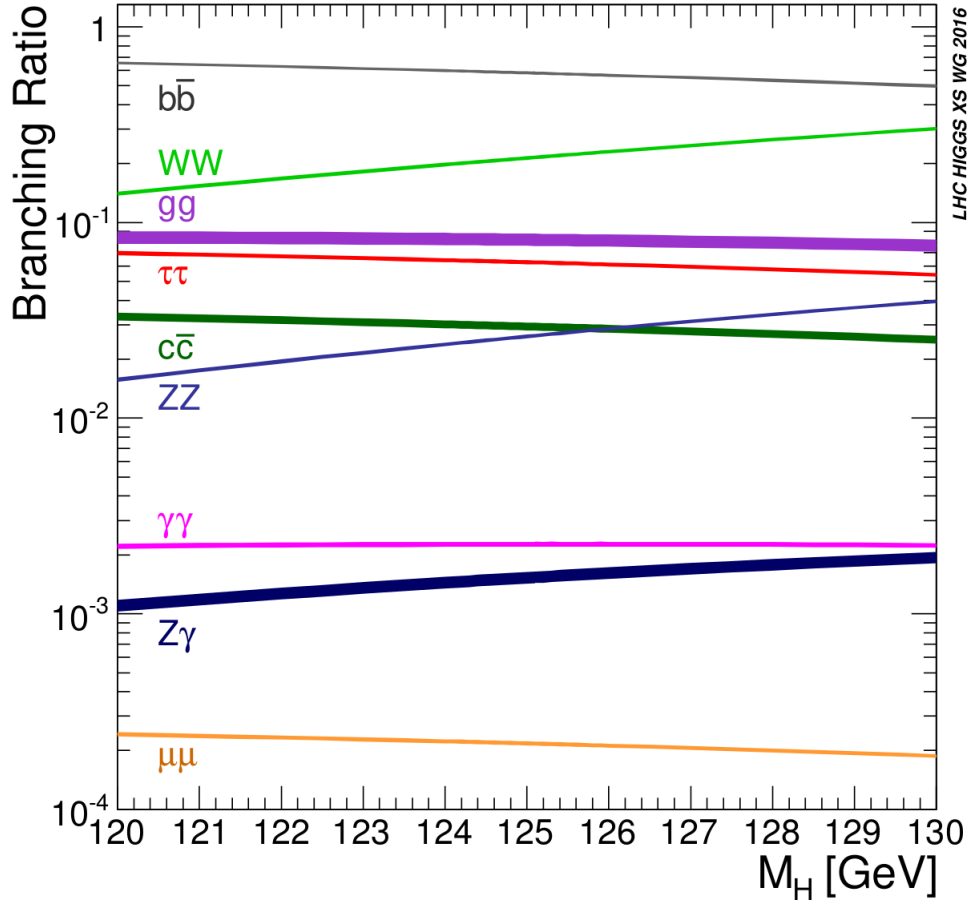


Рисунок 1 – Относительные вероятности основных распадов бозона Хиггса в Стандартной Модели при массе близкой к 125 ГэВ [7]

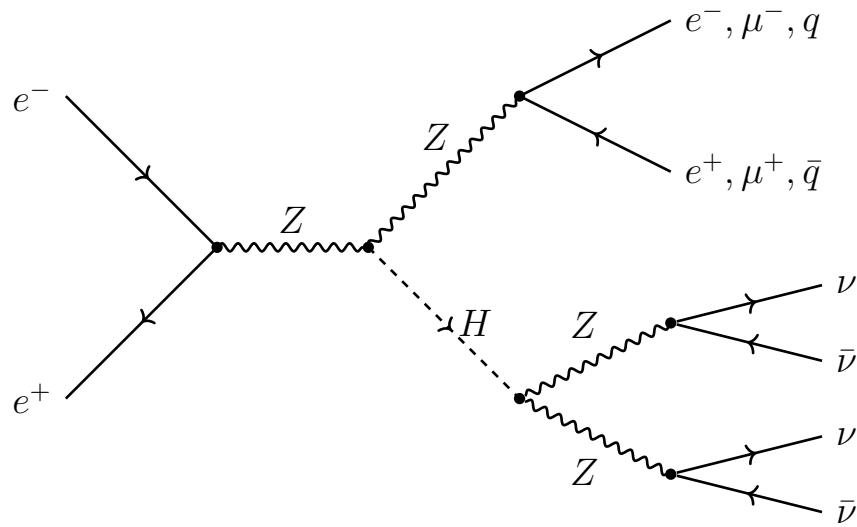


Рисунок 2 – Рождение бозона Хиггса в ассоциации с Z -бозоном и последующий распад $H \rightarrow ZZ^* \rightarrow 4\nu$

1 Обзор эксперимента CEPC

Круговой электрон-позитронный коллайдер (CEPC, Circular Electron Positron Collider) [8] — это крупный международный проект в области физики высоких энергий, инициированный китайским научным сообществом при широком участии международных партнёров. CEPC представляет собой двухкольцевой ускоритель (схематическое изображение на рисунке 1.1), в котором пучки электронов и позитронов циркулируют в противоположных направлениях по независимым трубам, а регистрация столкновений осуществляется детекторами, расположенными в двух точках взаимодействия (Interaction Points, IPs). Частицы рождаются и разгоняются в линейном ускорителе, затем попадают в кольцо бустера, где ускоряются до рабочей энергии, после чего инжектируются в основное кольцо коллайдера. Планируется разместить установку в подземном туннеле протяжённостью около 100 км на территории Китая.

Основная научная мотивация проекта заключается в создании “фабрики бозона Хиггса” — установки, оптимизированной для прецизионных измерений свойств хиггсовского бозона и электрослабых взаимодействий. По сравнению с адронными машинами, лептонные коллайдеры обладают значительно более чистой экспериментальной средой и точно известными начальными состояниями пучков. Это позволяет использовать метод массы отдачи (подробнее см. раздел 3.3), при котором инвариантная масса бозона Хиггса восстанавливается исключительно по измеренным параметрам сопровождающего Z -бозона, без необходимости реконструкции продуктов распада самого Хиггса. Благодаря такому подходу сечение рождения H и его свойства могут быть измерены модельно-независимым образом с точностью до процента и менее, что важно для поиска редких и невидимых распадов, обсуждавшихся ранее.

В настоящее время проект успешно прошёл стадии концептуального (CDR) [10; 11] и технического проектирования (TDR, завершён в 2023 г.) [9]. На текущий момент реализуется этап инженерного проектирования (Engineering Design Report, EDR, 2024–2026 г.), в ходе которого завершается окончательный выбор площадки, продолжаются исследования ключевых технологий, изготавливаются прототипы узлов и создаётся

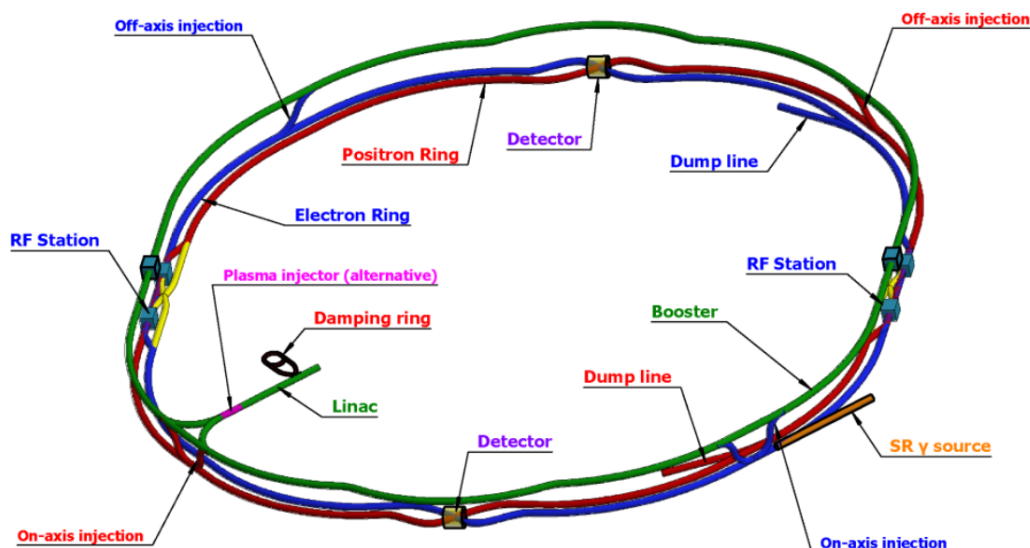


Рисунок 1.1 – Схематическое изображение коллайдера СЕРС. Коллайдер (Positron Ring, Electron Ring) и бустер (Booster) находятся в одном подземном туннеле, а линейный ускоритель (Linac) расположен на поверхности [9]

инфраструктура для промышленного производства компонентов. Согласно плану (рисунок 1.2), строительство коллайдера предполагается начать ориентировочно в 2027 году. Весь строительный цикл оценивается в 8 лет и его завершение планируется к 2035 году. Детальный обзор сценариев эксплуатации, конструкции детектора и программного обеспечения представлен в следующих подразделах.

1.1 Физическая программа и сценарии работы СЕРС

СЕРС спроектирован для работы в широком диапазоне энергий в системе центра масс ($\sqrt{s}$), охватывающем все ключевые пороги рождения интересных частиц: от Z -пика при 91 ГэВ и порога рождения W^+W^- -пар при 160 ГэВ до резонансного рождения бозона Хиггса в ассоциации с Z -бозоном при 240 ГэВ и порога рождения $t\bar{t}$ -пар при 340–360 ГэВ. Для ускорения ввода в эксплуатацию проектом предусмотрена двухэтапная стратегия — базовая и модернизированная, различающиеся максимальной мощностью синхротронного излучения (СИ) и конфигурацией точек взаимодействия.

Базовый сценарий предполагает работу с одной точкой взаимодействия (IP) и мощностью синхротронного излучения 30 МВт, в его рамках коллайдер начнёт работу при $\sqrt{s} = 240$ ГэВ в качестве фабрики бозонов Хиггса. В

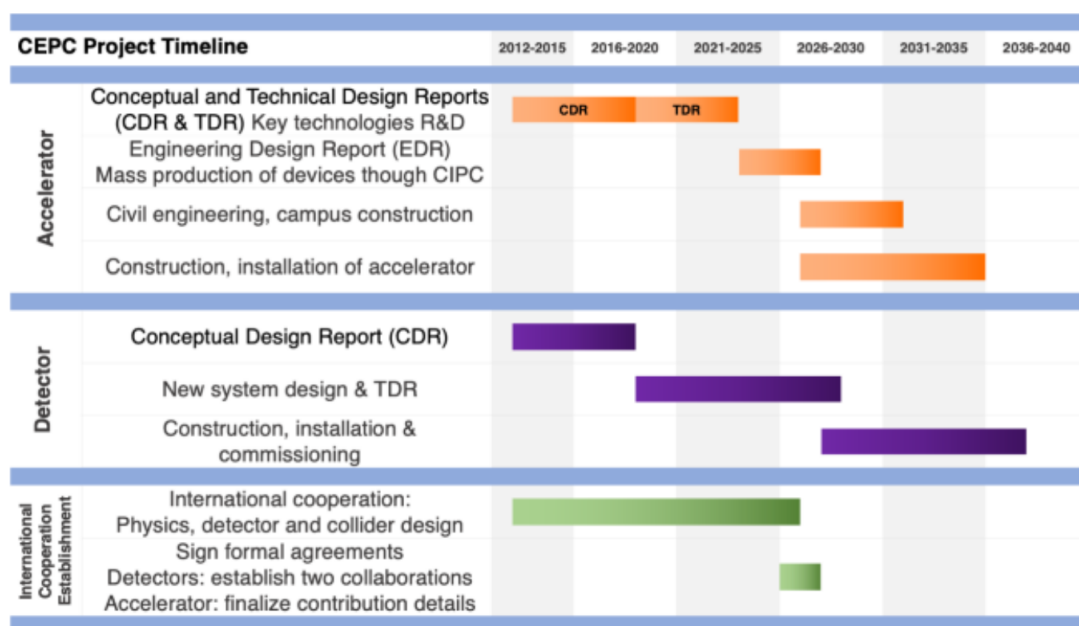


Рисунок 1.2 – Хронология проекта СЕРС с момента его первого предложения в 2012 году, включая разработку ускорителя, детектора и международное сотрудничество. Начало физических экспериментов ожидается в 2037 году [12]

течение 15 лет планируется накопить интегральную светимость 10 аб^{-1} , достаточную для регистрации около $2,0 \times 10^6$ событий ZH . Данный этап будет дополнен короткими периодами работы на Z -пике, необходимыми для калибровки детектора и изучения фоновых процессов. Впоследствии предусмотрен четырёхлетний режим “Супер- Z -фабрики” (мощность СИ 12,1 МВт) для получения $5,6 \times 10^{11}$ Z -бозонов, а также годовой этап сканирования порога рождения пар W^+W^- (155–170 ГэВ), что позволит измерить массу W -бозона с точностью $\sim 1 \text{ МэВ}$. Мгновенная светимость на Z -пике в базовом режиме составит $\sim 26 \times 10^{34} \text{ см}^{-2}\text{с}^{-1}$.

Модернизированный сценарий предполагает установку двух детекторов в точках взаимодействия и повышение мощности синхротронного излучения до 50 МВт. Это позволит сократить хиггсовскую программу до 10 лет, увеличив накопленную статистику до $4,3 \times 10^6$ бозонов Хиггса (интегральная светимость $21,6 \text{ аб}^{-1}$). Работа в режиме Z -фабрики при поле соленоида 2 Тл обеспечит сбор до $4,1 \times 10^{12}$ Z -бозонов за два года. Опциональное продление программы до энергий 340–360 ГэВ позволит набрать $0,6 \times 10^6$ событий рождения $t\bar{t}$ -пар за пятилетний период и измерить массу топ-кварка с точностью порядка 10 МэВ.

Ожидается, что точность измерения констант связи бозона Хиггса достигнет уровня 1% и ниже, а электрослабые параметры, такие как массы W - и Z -бозонов, будут определены с неопределённостью на уровне МэВ. В сочетании с чистой экспериментальной средой и методикой массы отдачи это позволит существенно превзойти текущие ограничения LHC и установить наиболее строгие верхние пределы на вероятность распада $BR(H \rightarrow \text{invisible})$, а также провести его модельно-независимое измерение. Реализация столь высоких требований к точности и статистике напрямую определяет архитектуру детекторной системы, концепция которой рассматривается в следующем разделе.

1.2 Базовый детектор CEPC

Базовый детектор CEPC представляет собой комплексную установку, оптимизированную для прецизионных исследований бозона Хиггса, электрослабых взаимодействий и поиска новой физики. В его конструкции широко применяются передовые технологии, разработанные за последние годы, при этом некоторые компоненты могут быть модернизированы в будущем. Характеристики детектора предполагают применение метода потока частиц (Particle Flow Algorithm, PFA) [13; 14] при реконструкции событий. Это подход, который объединяет информацию от трековых и калориметрических систем для реконструкции частиц с высокой точностью. Основная идея PFA заключается в измерении энергии каждой частицы внутри струи с использованием оптимальной подсистемы детектора: импульс заряженных частиц измеряется по кривизне треков в магнитном поле, энергия фотонов измеряется в электромагнитном калориметре, а энергия нейтральных адронов — в адронном калориметре.

Базовый детектор (рис. 1.3) является цилиндрической установкой, которая состоит из нескольких концентрических подсистем. От внутренней к внешней оболочке детектор включает: кремниевый пиксельный вершинный детектор (VTX), внутренний кремниевый трекер (ITK), времяпроекционную камеру (TPC), внешний кремниевый трекер (OTK), кристаллический электромагнитный калориметр (ECAL), адронный калориметр (HCAL) на основе стальных пластин и стеклянного сцинтиллятора, сверхпроводящий

соленоид с магнитным полем 3 Тл (Magnet), ярмо возврата магнитного потока с встроенным мюонным детектором (Yoke), а также специализированный калориметр для мониторинга светимости (LumiCal), установленный рядом с вакуумной камерой пучка (Beam pipe).

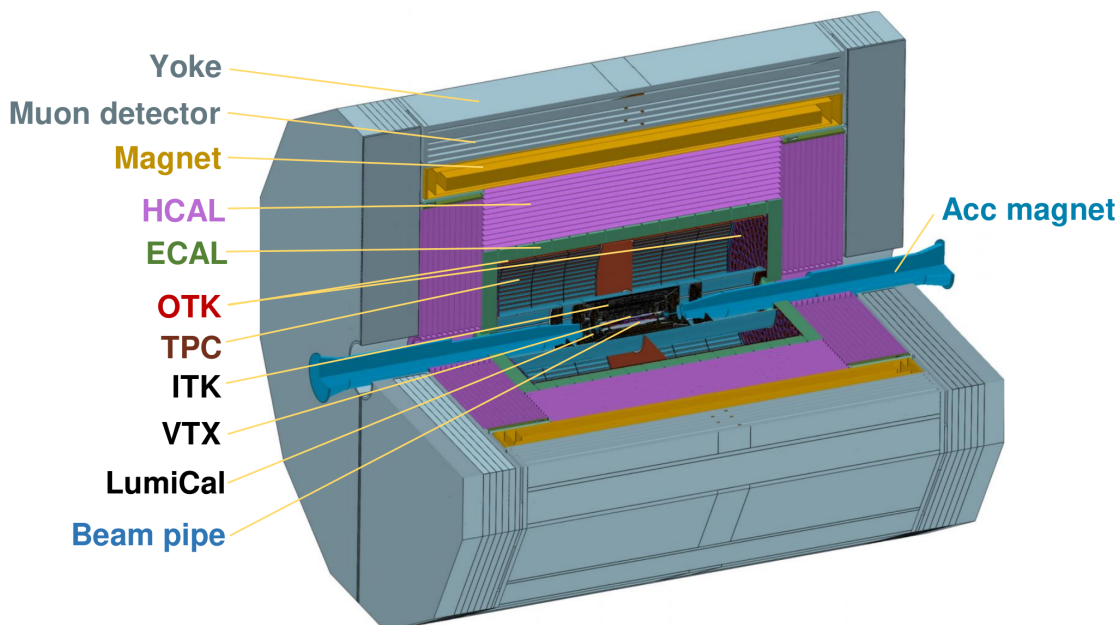


Рисунок 1.3 – Схематическое изображение конструкции базового детектора, демонстрирующее его основные подсистемы [12]; пояснения в тексте

1.2.1 Интерфейс сопряжения ускорителя и детектора

Интерфейс Machine Detector Interface (MDI) обеспечивает связь между коллайдером и детектором. Его центральным элементом является сверхтонкая бериллиевая вакуумная камера пучка, спроектированная для минимизации эффектов многократного рассеяния частиц. Для измерения светимости используются детекторы LumiCal, расположенные на расстоянии ± 560 мм от точки взаимодействия. Они состоят из кремниевых слоев на базе технологии AC-LGAD и кристаллов LYSO, что позволяет одновременно отслеживать профиль пучка и измерять интегральную светимость.

1.2.2 Трековая система

Трековая система является важнейшей частью детектора и состоит из четырёх основных подсистем. Она должна обеспечивать высокое разрешение по импульсу трека, необходимое для точных измерений свойств бозона

Хиггса.

Кремниевый пиксельный вершинный детектор (VTX) расположен максимально близко к точке взаимодействия и играет ключевую роль в идентификации таких частиц, как B -мезоны. Его компактная геометрия включает шесть концентрических слоёв пиксельных сенсоров, обеспечивающих трёхмерный трекинг с высоким пространственным разрешением. Эта точность важна для выделения вторичных вершин от распадов частиц с тяжёлыми кварками.

Внутренний кремниевый трекер (ITK) расположен вокруг VTX и включает три баррельных слоя кремниевых сенсоров, дополненных торцевыми дисками, установленными вдоль оси пучка. При проектировании ITK особое внимание уделено достижению высокого временного разрешения.

Времяпроекционная камера (TRC) является газовым детектором, который обеспечивает трёхмерную реконструкцию траекторий заряженных частиц. Он расположен в цилиндрическом объёме и разделен центральной катодной плоскостью на две дрейфовые области. Заряженные частицы ионизируют газовую смесь, а освобождённые электроны дрейфуют к плоскостям считывания под действием однородного электрического поля. В детекторе используется технология Micromegas, которая обеспечивает высокое пространственное разрешение и возможность измерения плотности ионизации вдоль трека, что позволяет разделять различные типы частиц.

Внешний кремниевый трекер (OTK) является самым внешним слоем трековой системы, объединяющим функции измерения импульса и детектирования по времени пролёта. Его геометрия включает один баррельный слой и два торцевых диска. В OTK применяются микростриповые сенсоры на базе технологии AC-LGAD, которые позволяют одновременно измерять координаты и время пролёта частиц. Данная возможность важна для идентификации и поисков долгоживущих частиц.

1.2.3 Калориметрическая система.

Калориметрическая система состоит из электромагнитного (ECAL) и адронного (HCAL) калориметров, специально оптимизированных для

применения алгоритма PFA. Метод PFA решает задачу улучшения энергетического разрешения за счёт идентификации и разделения откликов от различных частиц внутри струи благодаря высокой гранулярности подсистем детектора.

Электромагнитный калориметр (ECAL) базовой конструкции использует кристаллы германата висмута (BGO). Геометрия калориметра включает 18 продольных слоёв кристаллических стержней BGO, уложенных со взаимно ортогональной ориентацией в соседних слоях. Кристаллы BGO были выбраны благодаря их высокой плотности и хорошему световыходу, что позволяет достичь высокого энергетического разрешения для электронов и фотонов. Кремниевые фотоумножители (SiPM), сопряжённые с кристаллами, обеспечивают быстрое считывание сигнала. Ортогональная ориентация кристаллов и ультрамелкая гранулярность ECAL позволяют надёжно разделять перекрывающиеся ливни в условиях плотных струйных событий.

Адронный калориметр (HCAL) предназначен для измерения энергии адронов. Его геометрия включает баррельную и торцевую части из 48 слоев. Каждый слой сегментирован на ячейки для удовлетворения требований гранулярности алгоритма PFA. В качестве активного материала в HCAL используются стеклянные сцинтилляторы, имеющие высокую плотность и короткое время затухания свечения. Сцинтилляционный свет считывается кремниевыми фотоумножителями, при этом общее количество каналов составляет несколько миллионов.

1.2.4 Магнитная система.

Детектор оснащён сверхпроводящим соленоидным магнитом, способным работать в режимах 3 Тл или 2 Тл (для режима работы на Z -пике). Соленоид обеспечивает высокую однородность магнитного поля, что позволяет проводить точные измерения импульса частиц. Окружающее соленоид возвратное ярмо не только замыкает магнитный поток, но и служит конструктивной основой для размещения слоёв мюонного детектора. Модульная конструкция ярма позволяет осуществлять поэтапный монтаж и проводить будущие модернизации детектора.

1.2.5 Мюонный детектор.

Мюонный детектор является завершающим элементом системы идентификации частиц и предназначен для реконструкции мюонов и их надёжного отделения от адронов, проникающих сквозь калориметр. Он состоит из нескольких концентрических баррельных слоёв и торцевых секторов. Каждый слой собран на основе пластиковых сцинтилляторов, сигнал с которых считывается спектросмещающим волокном сопряженным с кремниевыми фотоумножителями. Треки мюонов проходят через несколько слоёв сцинтилляторов, интегрированных в структуру ярма, благодаря чему достигается высокая эффективность идентификации мюонов.

2 Программное обеспечение

2.1 Архитектура и базовые компоненты

Ядром программной инфраструктуры является программный пакет CEPCSW [15], который реализует трёхуровневую архитектуру (рисунок 2.1). Нижний уровень включает внешние библиотеки (ROOT, Geant4, DD4hep), средний уровень предоставляет инфраструктуру на базе пакета Gaudi [16] и интеграцию с интерфейсом Key4hep [17], верхний уровень содержит прикладные алгоритмы. Для описания детектора используется интерфейс DD4hep [18], который обеспечивает согласованность геометрии между моделированием и реконструкцией. Структура записи событий реализована на базе библиотеки EDM4hep [19], которая для хранения использует формат ROOT. Конвейер моделирования включает генерацию событий, моделирование в пакете Geant4 и оцифровку (рисунок 2.2). Для генерации процессов $e^+e^- \rightarrow ZH$ и фонов применяется пакет WHIZARD [20], адронизация обрабатывается пакетом Pythia [21].

Ключевым элементом реконструкции адронных струй и недостающей энергии является метод потока частиц (Particle Flow Algorithm, PFA). Для калориметра CEPSC ECAL на основе ортогональных кристаллических стержней BGO был разработан специализированный алгоритм в рамках пакета CyberPFA. Результатом работы алгоритма потока частиц является коллекция объектов-кандидатов Particle Flow Objects (PFO). Полный рабочий процесс алгоритма потока частиц включает несколько последовательных этапов:

- получение информации о реконструированных треках и сигналах в мюонной системе;
- кластеризация калориметрических хитов;
- ассоциация треков и кластеров в калориметрах;
- формирование гипотез о типе частиц (мюон, электрон, фотон, заряженный адрон, нейтральный адрон), вычисление четырехимпульсов.

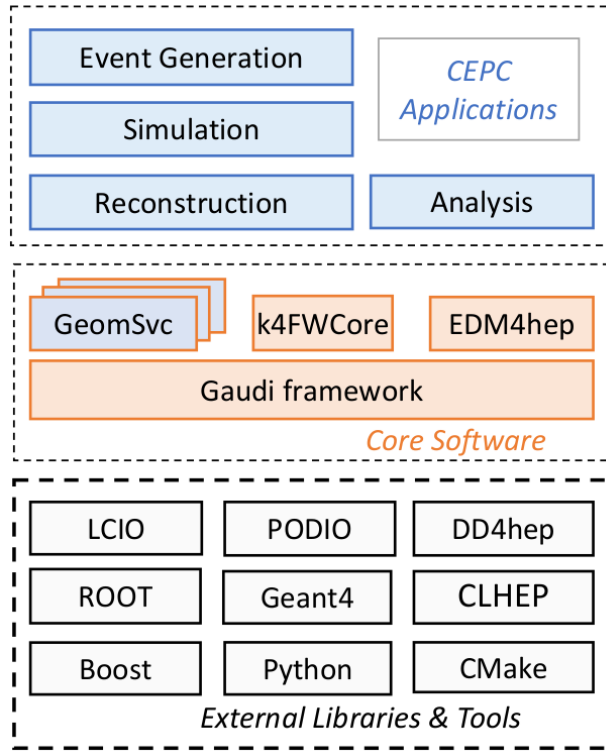


Рисунок 2.1 – Схематическое представление структуры программного обеспечения CEPCSW [12]

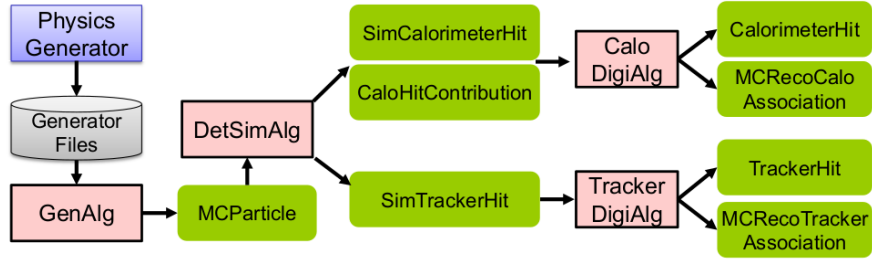


Рисунок 2.2 – Поток данных при моделировании [9]

2.2 Идентификация изолированных частиц

Выделение изолированных лептонов (электронов и мюонов) и фотонов из коллекции реконструированных объектов является критически важным этапом анализа для подавления фоновых процессов, таких как $Z \rightarrow l\bar{l}$ ($l = e, \mu$) или $W \rightarrow l\nu$ ($l = e, \mu$). В данной работе для базового детектора CEPС был разработан и адаптирован алгоритм идентификации изолированных частиц.

Для частицы-кандидата с импульсом $\vec{p}_i$ и другой частицы с импульсом $\vec{p}_j$ определяется косинус угла между трёхмерными импульсами частиц:

$$\cos \theta_{ij} = \frac{\vec{p}_i \cdot \vec{p}_j}{|\vec{p}_i| \cdot |\vec{p}_j|}. \quad (2.1)$$

Частица j считается находящейся внутри конуса вокруг частицы i , если выполняется условие:

$$\cos \theta_{ij} \geq \cos \theta_{\text{cone}}^{\text{cut}}, \quad (2.2)$$

где значение $\cos \theta_{\text{cone}}^{\text{cut}}$ ограничивает размеры конуса. Для каждой частицы-кандидата вычисляется суммарная энергия всех остальных частиц, попадающих в конус:

$$E_{\text{cone}} = \sum_{j \neq i, \cos \theta_{ij} \geq \cos \theta_{\text{cone}}^{\text{cut}}} E_j. \quad (2.3)$$

При проверке изоляции лептонов из суммы исключаются фотоны, что позволяет подавить влияние излучения в конечном состоянии (Final State Radiation, FSR).

Лептон (электрон или мюон) или фотон считается изолированным при одновременном выполнении двух групп условий. Первая группа задаёт допустимый диапазон энергии самого трека-кандидата: $E_{\text{track}}^{\text{min}} \leq E_{\ell} \leq E_{\text{track}}^{\text{max}}$. Вторая группа условий накладывает требования на энергию в конусе вокруг кандидата: $E_{\text{cone}}^{\text{min}} \leq E_{\text{cone}} \leq E_{\text{cone}}^{\text{max}}$.

Алгоритм поддерживает два режима проверки изоляции: прямоугольный (пороговый) критерий, описанный выше, и альтернативный полиномиальный критерий, в котором граница допустимой энергии в конусе задаётся квадратичной зависимостью от энергии трека. В рамках данной работы используется только прямоугольный критерий.

Для проверки корректности работы алгоритма и выбора оптимальных рабочих параметров было проведено исследование его эффективности на контрольном процессе $e^+e^- \rightarrow ZH$ ($Z \rightarrow \mu^+\mu^-$, $H \rightarrow \text{inclusive}$), гарантированно содержащем пару высокоэнергетических изолированных мюонов. На рисунке 2.3 показано распределение числа лептонов с энергией выше порога $E_{\text{track}}^{\text{min}} = 5$ ГэВ на событие. Как видно из распределения, подавляющее большинство событий содержит более двух лептонов, что соответствует ожидаемой топологии процесса $Z \rightarrow \mu^+\mu^-$.

Эффективность идентификации ε определялась как отношение числа событий, в которых алгоритм нашёл не менее двух изолированных лептонов

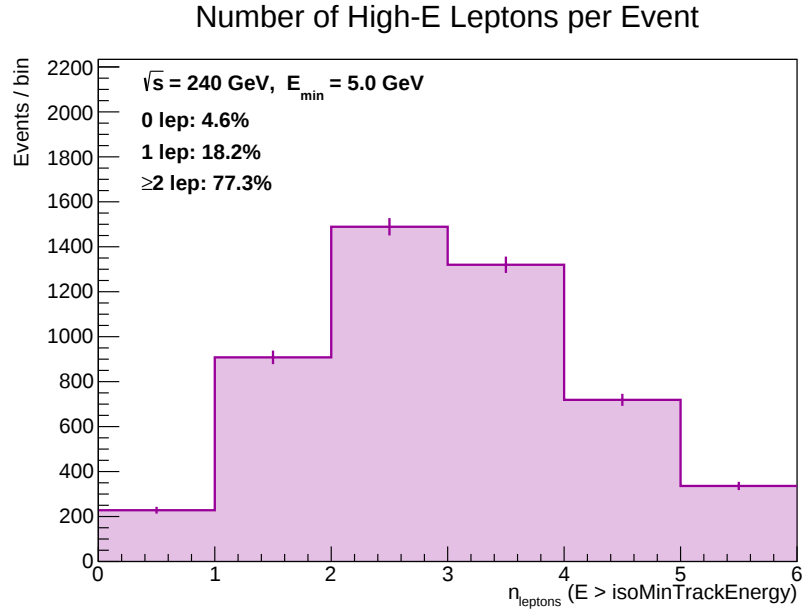


Рисунок 2.3 – Распределение числа реконструированных лептонов с энергией $E > E_{\text{track}}^{\min} = 5$ ГэВ на событие для процесса $e^+e^- \rightarrow ZH$ ($Z \rightarrow \mu^+\mu^-$, $H \rightarrow \text{inclusive}$)

$N(n_{\text{isolated}} \geq 2)$, к общему числу событий, содержащих хотя бы один лептон-кандидат с энергией выше порога $N(n_{E > E_{\text{track}}^{\min}} > 0)$:

$$\varepsilon = \frac{N(n_{\text{isolated}} \geq 2)}{N(n_{E > E_{\text{track}}^{\min}} > 0)} \times 100\%. \quad (2.4)$$

Был выполнен скан по трём ключевым параметрам критерия изоляции: минимальной энергии трека $E_{\text{track}}^{\min}$, максимальной допустимой энергии в конусе $E_{\text{cone}}^{\max}$ и параметру ширины конуса, заданному через $\cos \theta_{\text{cone}}$. Для каждого фиксированного значения $E_{\text{track}}^{\min}$ строились двумерные карты эффективности в плоскости $(\cos \theta_{\text{cone}}, E_{\text{cone}}^{\max})$. На рисунке 2.4 представлена двумерная карта эффективности алгоритма для $E_{\text{track}}^{\min} = 5.0$ ГэВ. По оси абсцисс отложены значения $\cos \theta_{\text{cone}}$ (с соответствующими углами в градусах), по оси ординат — значения $E_{\text{cone}}^{\max}$. Цветом обозначена эффективность в процентах.

На основании результатов сканирования были выбраны финальные значения параметров критерия изоляции, приведённые в таблице 2.1. Данные параметры обеспечивают высокую эффективность и используются во всех последующих этапах анализа.

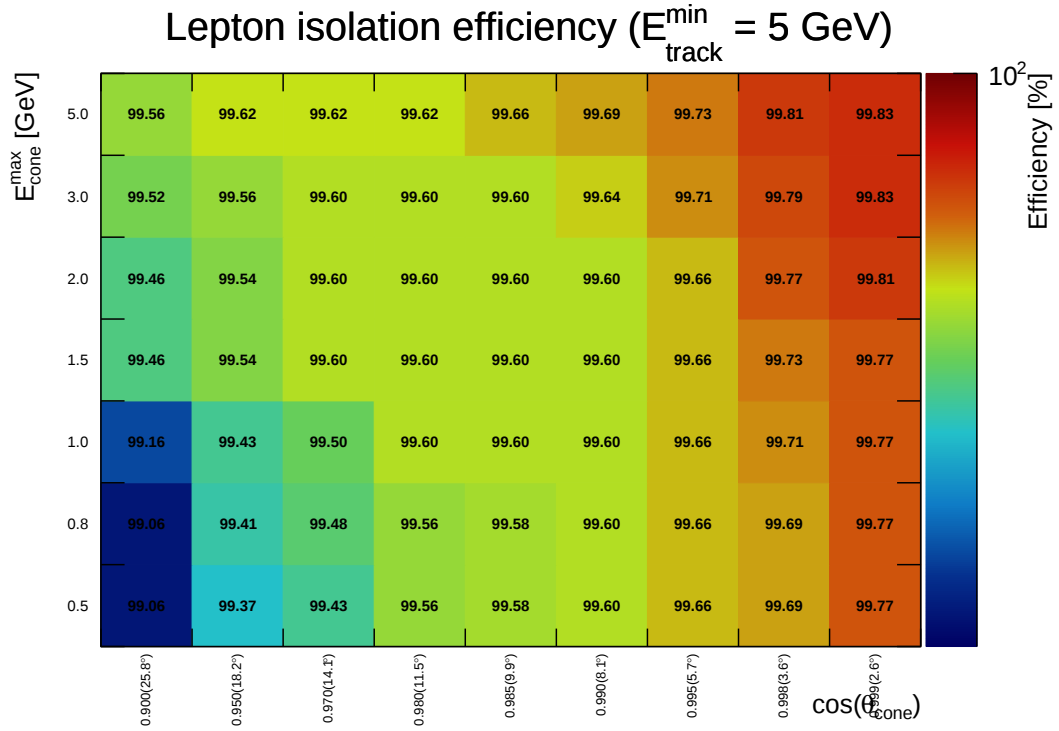


Рисунок 2.4 – Двумерная карта эффективности алгоритма изоляции лептонов в процессе $e^+e^- \rightarrow ZH$ ($Z \rightarrow \mu^+\mu^-$, $H \rightarrow \text{inclusive}$) для $E_{\text{track}}^{\min} = 5.0 \text{ GeV}$

Таблица 2.1 – Параметры алгоритма изоляции лептонов, выбранные по результатам исследования эффективности

Параметр	Обозначение	Значение
Минимальная энергия трека	$E_{\text{track}}^{\min}$	5.0 ГэВ
Минимальная энергия в конусе	$E_{\text{cone}}^{\min}$	0.0 ГэВ
Максимальная энергия в конусе	$E_{\text{cone}}^{\max}$	2.0 ГэВ
Размер конуса изоляции	$\cos \theta_{\text{cone}}^{\text{cut}}$	0.985 ($\theta_{\text{cone}} \approx 9.9^\circ$)

2.3 Пакет FastJet для анализа струй

Для поиска и анализа адронных струй (джетов) используется программный пакет FastJet [22]. Это С++ библиотека, которая предоставляет набор инструментов для кластеризации частиц в струи и исследования их внутренней структуры. Она включает эффективные реализации наиболее распространённых алгоритмов последовательной комбинации (k_t , anti- k_t , Cambridge/Aachen) как для pp -, так и для e^+e^- -столкновений, а также обеспечивает к ним унифицированный доступ через механизм плагинов.

Ключевой особенностью FastJet является оптимизация алгоритмов,

которая позволяет снизить асимптотическую сложность кластеризации с $\mathcal{O}(N^3)$ до $\mathcal{O}(N \ln N)$ для N входных частиц. Это делает пакет пригодным для обработки больших объёмов данных, характерных для современных экспериментов. Помимо оптимизации алгоритмов, пакет FastJet предоставляет средства для работы с подструктурой струй: вычисления инвариантных масс, угловых параметров, числа структурных элементов, энергетических корреляционных функций. Расширяемая архитектура позволяет интегрировать и адаптировать пользовательские алгоритмы под специфические задачи эксперимента.

В рамках данной работы FastJet используется на этапе реконструкции после формирования коллекции Particle Flow Objects. Для поиска адронных струй (джетов) был использован обобщённый алгоритм k_t для e^+e^- -столкновений. Он работал в инклюзивном режиме кластеризации. Данный алгоритм поиска джетов определяет расстояния между частицами следующим образом:

$$d_{ij} = \min(E_i^{2p}, E_j^{2p}) \frac{1 - \cos \theta_{ij}}{1 - \cos R}, \quad d_{iB} = E_i^{2p}, \quad (2.5)$$

где E_i — энергия i -й частицы, θ_{ij} — угол между частицами i и j , R — параметр радиуса джета, а p — внутренний параметр. При $p = 1$ алгоритм воспроизводит поведение классического k_t -алгоритма.

В инклюзивном режиме все реконструированные частицы (PFO), полученные после работы алгоритма CyberPFA, подаются на вход кластеризатора. Алгоритм последовательно объединяет пары частиц с минимальным расстоянием d_{ij} , пока не останется только набор «внутренних» джетов, для которых $d_{iB} < d_{jk}$ для всех оставшихся пар. В рамках анализа были выбраны параметры: $R = 0.5$, $p = 1.0$.

3 Анализ

3.1 Моделирование

Все сигнальные и фоновые процессы, рассматриваемые в данной работе, были смоделированы рабочей группой CEPС. Файлы в формате ROOT с результатами реконструкции доступны на серверах коллаборации.

3.1.1 Сигнальные и фоновые процессы

Основной сигнальный канал $e^+e^- \rightarrow ZH$ ($Z \rightarrow q\bar{q}$, $H \rightarrow \text{invisible}$) представлен двумя типами файлов реконструкции:

- qqHinv: чистый сигнальный процесс, в котором бозон Хиггса распадается исключительно по каналу $H \rightarrow ZZ^* \rightarrow 4\nu$;
- qqHX: процесс с инклюзивным распадом бозона Хиггса $H \rightarrow \text{inclusive}$ при фиксированном распаде Z -бозона в кварковую пару $Z \rightarrow q\bar{q}$ (рисунок 3.1). Данный файл реконструкции содержит как сигнальные события, так и фоновые конфигурации, возникающие из других каналов распада бозона Хиггса.

Инклюзивные распады бозона Хиггса (рисунок 1 и таблица 1) означают, что в конечном состоянии возможно различное количество джетов и изолированных лептонов. Топологии конечных состояний для семпла qqHX систематизированы в таблице 3.1. Анализ относительных вероятностей распадов показывает, что 85% событий содержат 4 джета, порядка 60% событий содержат 4 джета и ни одного изолированного лептона. Эти фоновые конфигурации будут учтены в последующем анализе в разделе 3.4.

Для реалистичной оценки чувствительности эксперимента необходимо учитывать основные фоновые процессы, способные имитировать сигнальную сигнатуру. В данной работе рассмотрены фоновые каналы, перечисленные в таблице 3.2.

Таблица 3.1 – Топологии конечных состояний при инклюзивном распаде бозона Хиггса в процессе $e^+e^- \rightarrow ZH$ ($Z \rightarrow q\bar{q}$, $H \rightarrow \text{inclusive}$)

Топология конечного состояния	Канал распада бозона Хиггса
2 джета + 0, 2 или 4 изолированных лептона + недостающая энергия	$H \rightarrow ZZ^* \rightarrow 4\nu$ (сигнальный)
	$H \rightarrow WW \rightarrow l\nu_l l\nu_l$
	$H \rightarrow \tau\tau \rightarrow l\nu_\tau \nu_l l\nu_\tau \nu_l$
	$H \rightarrow ZZ^* \rightarrow ll ll$
3 джета + 1 изолированный лептон + недостающая энергия	$H \rightarrow \tau\tau \rightarrow \text{hadr} l\nu_\tau \nu_l$
4 джета + 0, 1 или 2 изолированных лептона + [недостающая энергия]	$H \rightarrow \tau\tau \rightarrow \text{hadr hadr}$
	$H \rightarrow b\bar{b}, c\bar{c}, gg$
	$H \rightarrow WW \rightarrow qq l\nu_l$
	$H \rightarrow ZZ \rightarrow qq ll$
6 джетов	$H \rightarrow WW \rightarrow qq qq$
	$H \rightarrow ZZ \rightarrow qq qq$

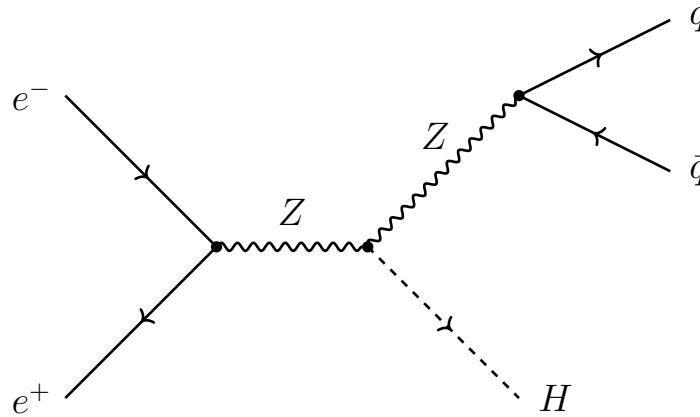


Рисунок 3.1 – Процесс с инклюзивным распадом бозона Хиггса $e^+e^- \rightarrow ZH$ ($Z \rightarrow q\bar{q}$, $H \rightarrow \text{inclusive}$)

3.1.2 Статистика реконструированных событий

В таблице 3.3 приведена сводная информация по всем обработанным процессам: сечение рождения, ожидаемое число событий N_{exp} при интегральной светимости $\mathcal{L} = 5050 \text{ фб}^{-1}$, число успешно реконструированных событий N_{ana} , а также статистический вес $W = N_{\text{exp}}/N_{\text{ana}}$, используемый для нормировки гистограмм при построении распределений и аппроксимации. Все процессы были обработаны на вычислительном кластере HTCondor [23]. Общее число анализируемых событий составило около 3.9 млн, что обеспечивает достаточную статистику для оценки эффективности отборов и построения распределений массы

Таблица 3.2 – Основные фоновые процессы, рассматриваемые в анализе

Обозначение процесса	Конечное состояние
qq	$e^+e^- \rightarrow q\bar{q}$
4f_sw_sl0qq	$e^+e^- \rightarrow e\nu_e q\bar{q}$
4f_ww_h0cuxx	$e^+e^- \rightarrow W^+W^- \rightarrow q\bar{q}q\bar{q}$
4f_sze_sl0uu	$e^+e^- \rightarrow Ze^+e^- \rightarrow q\bar{q}e^+e^-$
4f_zz_h0dtdt	$e^+e^- \rightarrow ZZ \rightarrow q\bar{q}q\bar{q}$
4f_sznu_sl0nu_up	$e^+e^- \rightarrow Z\nu\bar{\nu} \rightarrow q\bar{q}\nu\bar{\nu}$
nunuHX	$e^+e^- \rightarrow \nu\bar{\nu}H(\text{inclusive})$

отдачи.

Таблица 3.3 – Статистика смоделированных и реконструированных процессов при $\sqrt{s} = 240$ ГэВ и $\mathcal{L} = 5050 \text{ фб}^{-1}$

Процесс	σ , фб	N_{exp}	N_{ana}	W
qqHinv (сигнальный)	0.143	724	135 000	0.0054
qqHX	143.39	724 120	892 750	0.8111
qq	50 105.35	253 032 018	492 524	513.7456
4f_sw_sl0qq	2 579	13 023 950	396 400	32.8556
4f_sze_sl0uu (dd)	325	1 641 250	398 800	4.1155
4f_sznu_sl0nu_up (down)	147	742 350	398 400	1.8633
4f_ww_h0cuxx	3 733	18 851 650	395 400	47.6774
4f_zz_h0dtdt	501	2 530 050	394 200	6.4182
nunuHX	48.96	247 248	836 920	0.2954

3.2 Предварительные отборы

Для эффективного подавления фоновых процессов и выделения сигнальной топологии $e^+e^- \rightarrow ZH$ ($Z \rightarrow q\bar{q}$, $H \rightarrow \text{invisible}$) в анализе применяется двухэтапная стратегия отбора событий: предварительные отборы и основные отборы. Предварительные отборы нацелены на исключение событий с явно несовместимой с сигналом топологией, что позволяет существенно сократить объём данных перед применением более тонких критериев. Все предотборы применяются последовательно в порядке, описанном ниже (число событий для всех каналов после каждого предотбора представлено в сводной таблице 3.4):

— изолированные лептоны: событие отбрасывается при наличии

хотя бы одного изолированного лептона. Изолированность определяется с помощью алгоритма, который был описан в разделе 2.2;

- высокоэнергетические фотоны: событие отбрасывается при наличии фотона с $E_\gamma > 30.0$ ГэВ (рисунок 3.2a);

- изолированные фотоны: событие отбрасывается при наличии фотона, который удовлетворяет критериям изоляции. Для определения изолированных фотонов также используется алгоритм, который был описан в разделе 2.2;

- ровно два джета: событие отбрасывается, если алгоритм кластеризации, описанный в разделе 2.3, реконструирует число инклюзивных джетов, отличное от двух (рисунок 3.2b);

- число частиц в джетах: каждый из двух джетов должен содержать не менее $N_{\text{const}} \geq 6$ частиц (рисунок 3.2c), что позволяет исключить события с адронными распадами τ лептонов.

3.3 Основные отборы

События, прошедшие предотборы, характеризуются наличием ровно двух адронных джетов без изолированных лептонов и высокоэнергетических фотонов. Для дальнейшего выделения сигнального процесса применяется последовательность кинематических отборов, основанных на восстановленных переменных и перечисленных ниже:

- недостающая поперечная энергия (MET): событие сохраняется, если величина $20 \text{ ГэВ} < \text{MET} < 60 \text{ ГэВ}$ (рисунок 3.3a). MET рассчитывается как модуль векторной суммы поперечных импульсов двух реконструированных джетов, взятой с обратным знаком:

$$\text{MET} = \left| \vec{p}_T^{jj} \right| = \sqrt{(p_x^{j1} + p_x^{j2})^2 + (p_y^{j1} + p_y^{j2})^2}; \quad (3.1)$$

- азимутальный угол между джетами ($\Delta\phi$): событие сохраняется, если $\Delta\phi_{jj} < 3.0$ радиан (рисунок 3.3b). Азимутальный угол между джетами вычисляется как разность их азимутальных углов ϕ :

$$\Delta\phi_{jj} = |\phi_1 - \phi_2|; \quad (3.2)$$

Таблица 3.4 – Статистика прохождения предварительных отборов для всех процессов (в скобках указана эффективность отбора относительно предыдущего этапа)

Процесс	Всего событий	Изолированный лептон	Фотон ($E > 30$ ГэВ)	Изолированный фотон	2 джета	≥ 6 частиц
qqHinv	135 000	135 000 (100.00%)	135 000 (100.00%)	135 000 (100.00%)	87 226 (64.61%)	86 249 (98.88%)
qqHX	892 750	695 139 (77.86%)	663 821 (95.49%)	631 181 (95.08%)	2 940 (0.47%)	2 873 (97.72%)
qq	492 524	466 139 (94.64%)	356 224 (76.42%)	343 662 (96.47%)	211 560 (61.56%)	209 739 (99.14%)
4f_sw_sl0qq	396 400	188 656 (47.59%)	164 724 (87.31%)	156 695 (95.13%)	11 781 (7.52%)	8 914 (75.66%)
4f_size_sl0uu	398 800	172 810 (43.33%)	151 578 (87.71%)	141 144 (93.12%)	43 923 (31.12%)	20 508 (46.69%)
4f_sznu_sl0nu_up	398 400	379 470 (95.25%)	360 827 (95.09%)	344 329 (95.43%)	209 670 (60.89%)	206 601 (98.54%)
4f_ww_h0cuxx	395 400	358 974 (90.79%)	317 636 (88.48%)	300 176 (94.50%)	536 (0.18%)	534 (99.63%)
4f_zz_h0tdtdt	394 200	345 571 (87.66%)	315 939 (91.43%)	299 134 (94.68%)	1 520 (0.51%)	1 514 (99.61%)
nunuHX	836 920	836 920 (100.00%)	836 920 (100.00%)	836 920 (100.00%)	399 670 (47.75%)	343 934 (86.05%)

— полярный угол системы двух джетов ($\cos \theta_Z$): событие сохраняется, если $|\cos \theta_Z| < 0.7$ (рисунок 3.3с), рассматривается только центральная часть детектора. Полярный угол θ_Z системы двух джетов вычисляется через её трёхмерный импульс $\vec{p}_{jj}$:

$$\cos \theta_Z = \frac{p_z^{jj}}{|\vec{p}_{jj}|}, \quad \text{где} \quad \vec{p}_{jj} = \vec{p}_{j1} + \vec{p}_{j2}; \quad (3.3)$$

— инвариантная масса пары реконструированных джетов (M_{jj}): событие сохраняется, если $75 \text{ ГэВ} < M_{jj} < 105 \text{ ГэВ}$ (рисунок 3.3d). Инвариантная масса системы двух джетов вычисляется через их четырёхимпульсы:

$$M_{jj} = \sqrt{(E_{j1} + E_{j2})^2 - |\vec{p}_{j1} + \vec{p}_{j2}|^2}; \quad (3.4)$$

— модуль недостающего трёхимпульса ($|\vec{P}_{\text{miss}}|$): событие сохраняется, если $20 \text{ ГэВ} < |\vec{P}_{\text{miss}}| < 80 \text{ ГэВ}$ (рисунок 3.3е). Вектор недостающего импульса вычисляется как противоположный сумме импульсов всех реконструированных частиц (PFO) в событии:

$$\vec{P}_{\text{miss}} = - \sum_k \vec{p}_k, \quad |\vec{P}_{\text{miss}}| = \sqrt{P_{\text{miss},x}^2 + P_{\text{miss},y}^2 + P_{\text{miss},z}^2}. \quad (3.5)$$

В таблице 3.5 приведена сводная статистика прохождения каждого этапа основных отборов для всех рассматриваемых процессов. Числа в ячейках показывают абсолютное количество событий, прошедших данный отбор, а в скобках показан процент от числа событий после предотборов (предыдущий столбец). Как видно из таблицы последовательность отборов эффективно подавляет фоновые каналы, а для сигнального процесса $qq\text{H}inv$ сохраняется около 34% событий. Это обеспечивает высокую эффективность выделения сигнала.

На рисунке 3.4 представлена суммарная гистограмма распределения массы отдачи M_{recoil} после применения полной последовательности отборов для всех рассматриваемых процессов, для которых осталась достаточная статистика. Все процессы на гистограмме нормированы на ожидаемое число событий при интегральной светимости $\mathcal{L} = 5050 \text{ фб}^{-1}$ с использованием весов

$W = N_{\text{exp}}/N_{\text{ana}}$, приведённых в таблице 3.3. Как видно из распределения, сигнальный канал формирует пик в области $M_{\text{recoil}} \approx 125$ ГэВ, что соответствует массе бозона Хиггса, в то время как фоновые процессы дают относительно гладкое распределение в данной области.

3.4 Сравнение файлов с чистым и инклюзивным сигналом

Как было описано в разделе 3.1, сигнальный процесс представлен двумя типами файлов: qqHinv и qqHX. Несмотря на то, что оба файла проходят идентичную последовательность отборов, описанную в разделах 3.2 и 3.3, их состав после применения всех критериев различается. Файл, который содержит процесс qqHinv представляет собой “чистый” сигнал, в котором вся недостающая энергия обусловлена исключительно четырьмя нейтрино от распада бозона Хиггса. В то же время, файл qqHX после отборов все равно содержит “загрязнённый” сигнал: помимо интересующего канала $H \rightarrow 4\nu$ в нём присутствуют события с другими распадами бозона Хиггса.

На рисунке 3.5 представлено сравнение распределений массы отдачи M_{recoil} для файлов qqHinv и qqHX после применения полной последовательности отборов. Для корректного сравнения формы распределений оба процесса были нормированы на одинаковое ожидаемое число сигнальных событий. В процессе qqHX с инклюзивным распадом Хиггса доля интересующего канала $H \rightarrow 4\nu$ составляет $\sim 0.1\%$ от всех распадов, что при 892 750 реконструированных событиях даёт ~ 893 ожидаемых сигнальных события. Данный процесс был взят с единичным весом ($W_{\text{qqHX}} = 1$), чтобы отобразить именно это количество сигнальных событий. Для чистого сигнального процесса qqHinv, содержащего 135 000 реконструированных событий исключительно канала, был рассчитан нормировочный вес $W_{\text{qqHinv}} = 893/135\,000 \approx 0.00661$, обеспечивающий одинаковую нормировку по числу ожидаемых сигнальных событий в обоих распределениях.

Как видно из распределения, оба процесса демонстрируют пик в области $M_{\text{recoil}} \approx 125$ ГэВ, соответствующий массе бозона Хиггса. Однако qqHX имеет более широкое распределение, что обусловлено вкладом фоновых каналов инклюзивного распада Хиггса. Эта особенность будет учтена при оценке

Таблица 3.5 – Статистика прохождения основных отборов для всех процессов (в скобках указана эффективность отбора относительно предыдущего этапа)

Процесс	После предотборов	MET	$\Delta\phi$	$ \cos\theta_Z $	M_{jj}	$ \vec{P}_{\text{miss}} $	Финал
qqHinv	86 249	75 201 (87.19%)	70 899 (94.28%)	56 103 (79.13%)	46 344 (82.61%)	46 336 (99.98%)	46 336 (34.32%)
qqHX	2 873	1 267 (44.10%)	1 030 (81.29%)	751 (72.91%)	513 (68.31%)	512 (99.81%)	512 (0.06%)
qq	209 739	7 635 (3.64%)	685 (8.97%)	55 (8.03%)	6 (10.91%)	6 (100.00%)	6 (0.00%)
4f_sw_sl0qq	8 914	5 538 (62.13%)	4 968 (89.71%)	2 071 (41.69%)	255 (12.31%)	249 (97.65%)	249 (0.06%)
4f_size_sl0uu	20 508	1 618 (7.89%)	1 524 (94.19%)	709 (46.52%)	1 (0.14%)	0 (0.00%)	0 (0.00%)
4f_sznu_sl0nu_up	206 601	117 915 (57.07%)	109 392 (92.77%)	42 814 (39.14%)	30 960 (72.31%)	30 043 (97.04%)	30 043 (7.54%)
4f_ww_h0cuxx	534	40 (7.49%)	21 (52.50%)	6 (28.57%)	1 (16.67%)	1 (100.00%)	1 (0.00%)
4f_zz_h0dtdt	1 514	102 (6.74%)	38 (37.25%)	15 (39.47%)	1 (6.67%)	1 (100.00%)	1 (0.00%)
nunuHX	343 934	281 352 (81.80%)	254 459 (90.44%)	197 593 (77.65%)	48 003 (24.29%)	47 925 (99.84%)	47 925 (5.73%)

верхнего предела в разделе 3.5.

3.5 Результаты аппроксимации

Для оценки потенциальной чувствительности анализа к сигнальному процессу выполнялась аппроксимация распределения массы отдачи M_{recoil} методом расширенного максимума правдоподобия (Extended Maximum Likelihood, EML) с использованием библиотеки RooFit [24]. Подгонка проводилось в диапазоне $100 \text{ ГэВ} < M_{\text{recoil}} < 155 \text{ ГэВ}$, охватывающем область пика сигнала и прилегающие фоновые области. Форма распределения сигнала задавалась непараметрическим шаблоном f_S , который был построен методом ядерной оценки плотности (Kernel Density Estimation) [25]. Фоновый шаблон f_B был построен с помощью полинома Чебышева четвертой степени. Шаблоны для сигнала и фона изображены на рисунке 3.6.

Для каждого процесса (сигнального $qq\text{Hinv}$ и совокупности фоновых) формировалась выборка, в которой каждому событию присваивался статистический вес W , приведённый в таблице 3.3. Псевдоданные для аппроксимации формировались путём объединения фоновых событий (с исходными весами) и сигнальных событий, веса которых дополнительно масштабировались на коэффициент μ , который варьировался от 0.0 до 5.0 с шагом 0.2. Этот коэффициент является множителем для сигнала относительно предсказания Стандартной Модели. Такая процедура позволяет оценить верхний предел на интенсивность сигнала, критерием оценки которого является достижение уровня значимости $p\text{-value} = 0.05$ при каком-то значении μ для проверяемых статистических гипотез:

— нулевая гипотеза H_0 (фон): аппроксимация псевдоданных моделью, содержащей только нормированный фоновый шаблон f_B . Свободным параметром является число фоновых событий n_B ;

— альтернативная гипотеза H_1 (сигнал + фон): аппроксимация полной моделью $n_S \cdot f_S(M_{\text{recoil}}) + n_B \cdot f_B(M_{\text{recoil}})$, где f_S и f_B — нормированные шаблоны сигнала и фона, а n_S и n_B — свободные параметры, представляющие число событий сигнала и фона.

Значимость оценивалась через отношение правдоподобий (Likelihood Ratio Test, LRT). Статистика теста определяется как $q_0(\mu) = -2 \ln(L_\mu/L_0) =$

$2 \cdot (\text{NLL}_0 - \text{NLL}_\mu)$, где NLL_0 и NLL_μ — значение отрицательного логарифма функции правдоподобия в минимуме для гипотезы H_0 и H_1 с заданным μ соответственно. По теореме Вилкса, при верной нулевой гипотезе q_0 асимптотически распределена как χ^2 с одной степенью свободы, что позволяет выразить значимость в единицах стандартного отклонения: $Z_{\text{LRT}} = \sqrt{q_0}$.

На рисунке 3.7 представлена зависимость p -value от μ . Найденным пороговым значением для доверительного интервала 95% является $\mu = 1.37$. На рисунке 3.8 представлен результат аппроксимации распределения массы отдачи для выборки, которая была построена с сигналом, умноженным на пороговое значение $\mu = 1.37$. Псевдоданные показаны точками с ошибками, рассчитанными как $\sqrt{\sum w_i^2}$ для каждого бина.

3.6 BDT-классификатор для подавления фоновых процессов

Основные отборы, которые были описаны в предыдущих разделах, позволяют хорошо подавить фоновые процессы, однако после них остается два доминирующих фона, которые имеют похожую на сигнал топологию: nunuHX и $4f_sznu_sl0nu_up$. Для их дополнительного подавления был применен метод градиентного бустинга, который реализован в библиотеке XGBoost [26]. В качестве признаков использовались следующие кинематические переменные:

- инвариантная масса пары джетов M_{jj} ;
- косинус полярного угла диджетной системы $\cos \theta_Z$;
- угловое расстояние между джетами ΔR ;
- косинус полярного угла каждого джета $\cos \theta_1, \cos \theta_2$;
- энергии джетов E_{j1}, E_{j2} ;
- недостающая поперечная энергия MET;
- энергия диджетной системы E_{jj} ;
- разность полярных углов джетов $\Delta \theta$;
- азимутальный угол между джетами $\Delta \phi$;
- модуль недостающего трёхимпульса $|\vec{P}_{\text{miss}}|$;
- косинус угла недостающего импульса $\cos \theta_{\text{miss}}$.

Для обучения использовалась выборка из событий, которые прошли через предварительные отборы. Она разделялась на аналитическую и обучающую выборки в соотношении 1 : 1. Обучающая выборка разделялась на тренировочную (70%) и тестовую (30%) части, на аналитической выборке производился инференс классификаторов. Оба классификатора обучались с одинаковыми гиперпараметрами:

- максимальная глубина дерева: 5;
- темп обучения: 0.05;
- доля событий для построения каждого дерева: 0.8;
- доля признаков для каждого дерева: 0.8;
- минимальный вес листа: 5;
- параметр гамма: 0.1;
- число деревьев: до 1000 с ранней остановкой через 50 раундов без улучшения на тестовой выборке.

Качество классификаторов оценивалось по площади под ROC-кривой (AUC), результаты обучения представлены в таблице 3.6. ROC кривые для классификаторов представлены на рисунке 3.9. Также был проанализирован вклад каждого входного признака, который оценивался с помощью метрики gain [26] (рисунок 3.10). Можно видеть, у обоих классификаторов есть три доминирующих признака, что позволяет предположить возможность объединения классификаторов в единый классификатор. Для выбора порога было построено распределение выхода классификаторов (рисунок 3.11). Были выбраны пороги 0.75 и 0.7 для qqHinvi vs nunuHX и qqHinvi vs 4f_sznu_sl0nu_up соответственно.

Таблица 3.6 – Характеристики обученных BDT-классификаторов

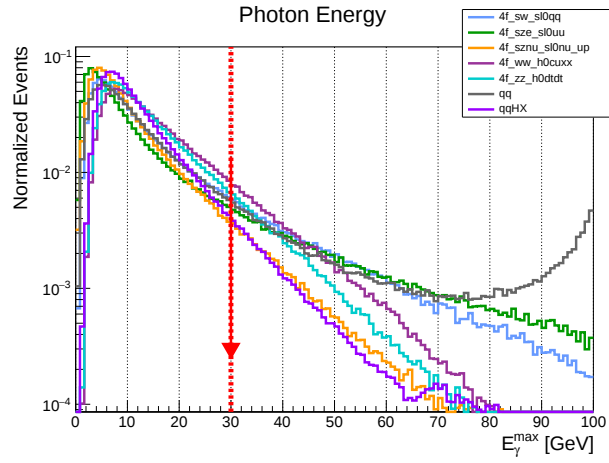
Классификатор	Число деревьев	AUC (train)	AUC (test)
qqHinvi vs nunuHX	721	0.9747	0.9558
qqHinvi vs 4f_sznu_sl0nu_up	287	0.9448	0.9302

Обученные модели были применены к аналитической выборке событий (после предотборов) с выбранными выше порогами. Для оценки эффективности классификаторов проведено сравнение с кинематическими отборами, результаты которого представлены в таблице 3.7. Можно видеть,

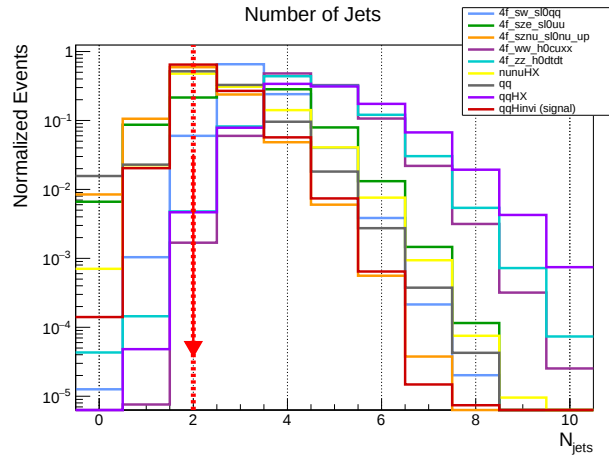
что фон $\mu\mu\mu HX$ с помощью BDT подавляется значительно лучше, чем с помощью кинематических отборов. Фон $4f_sznu_sl0nu_up$ имеет меньший выигрыш, но все же подавляется лучше кинематических отборов. Это позволяет сделать вывод, что использование BDT имеет потенциал по улучшению соотношения сигнала к фону в будущей работе. В текущем анализе дополнительное подавление с помощью BDT не реализовано.

Таблица 3.7 – Сравнение эффективности кинематических отборов и BDT-классификаторов (после предварительных отборов)

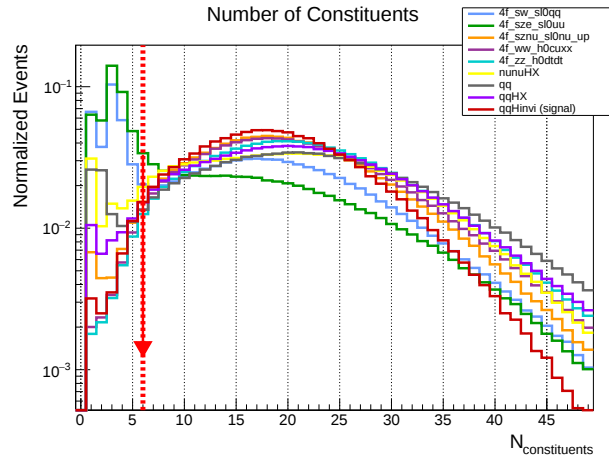
Процесс	Кинематические отборы	BDT-классификатор
$qqHinv$ (сигнал)	53.7%	81.3%/82.8%
$\mu\mu\mu HX$ (фон)	13.9%	7.0%
$4f_sznu_sl0nu_up$ (фон)	14.5%	12.1%



(a) Энергия реконструированного фотона E_γ

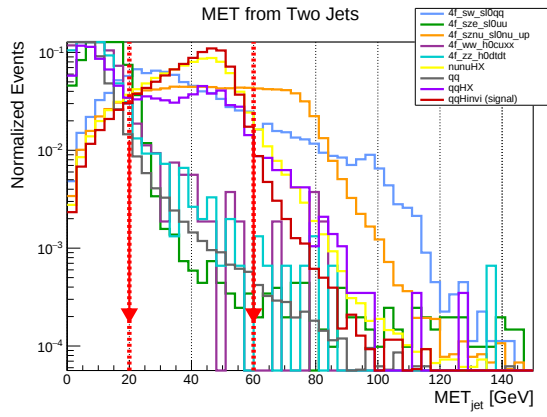


(b) Число реконструированных джетов n_{jets}

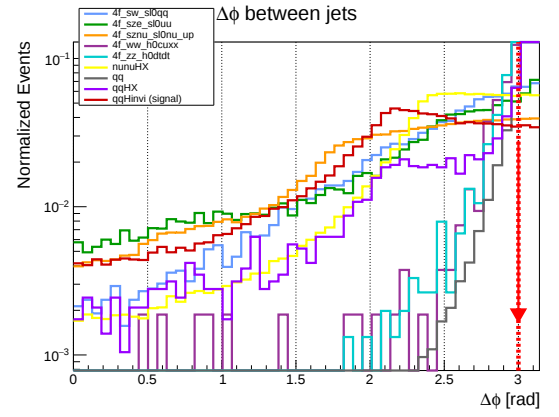


(c) Число частиц в джете N_{const}

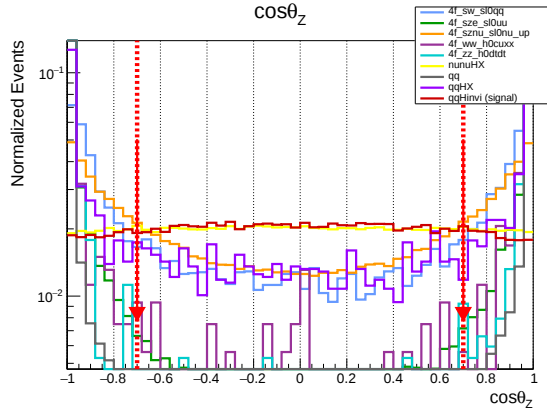
Рисунок 3.2 – Распределения переменных для всех процессов до предотборов, используемые для выбора пороговых значений предотборов. Красные вертикальные линии — пороговые значения



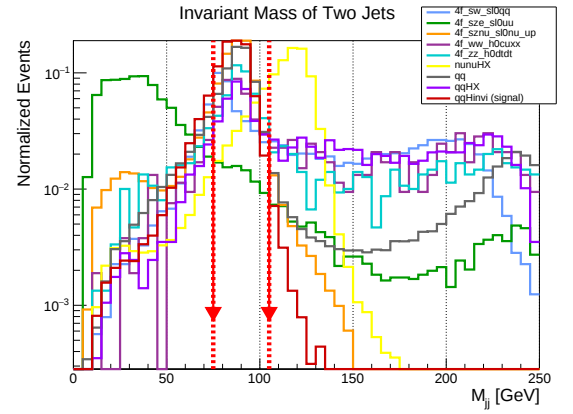
(a) Недостаточная поперечная энергия
MET



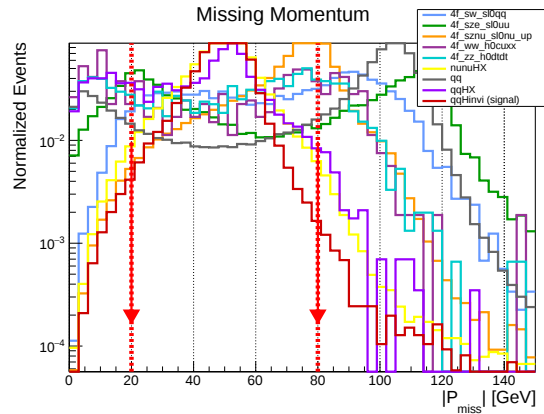
(b) Разница азимутальных углов двух
джетов $\Delta\phi$



(c) Косинус угла системы двух джетов
 $\cos\theta_z$



(d) Инвариантная масса двух джетов
 M_{jj}



(e) Модуль недостающего импульса
 $|\vec{P}_{\text{miss}}|$

Рисунок 3.3 – Распределения кинематических переменных для всех процессов после предотборов, используемые для выбора пороговых значений основных отборов. Красные вертикальные линии — пороговые значения

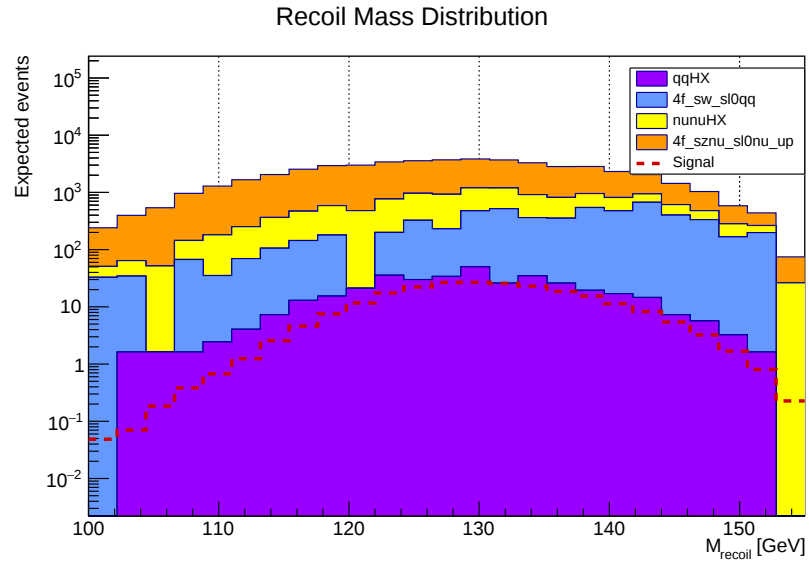


Рисунок 3.4 – Распределение массы отдачи M_{recoil} после применения всех отборов для фоновых процессов (заливка) и для сигнального процесса (красная пунктирная линия)

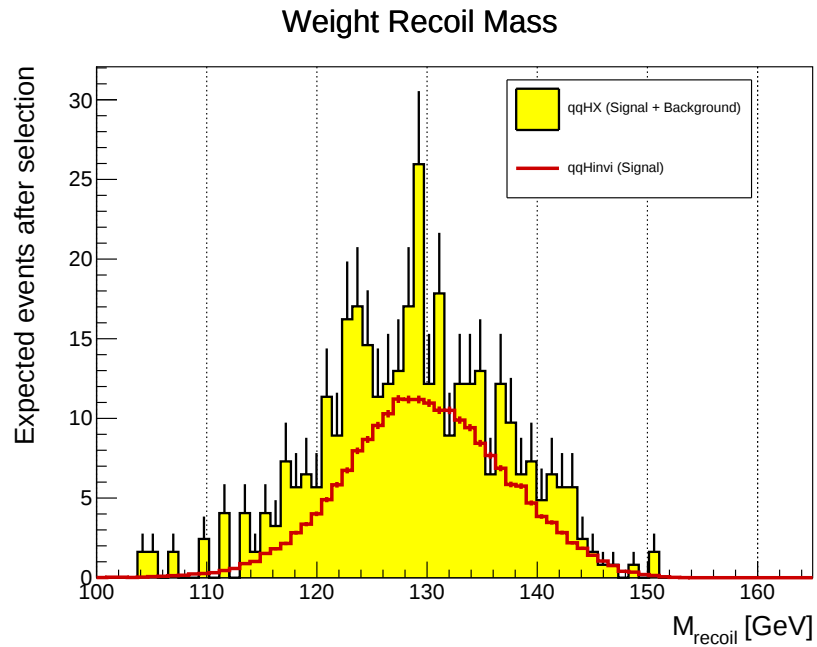
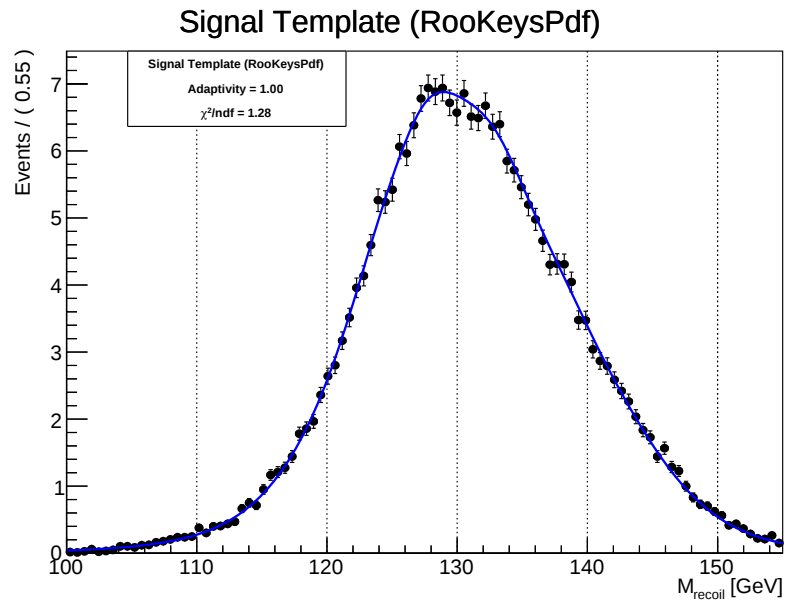
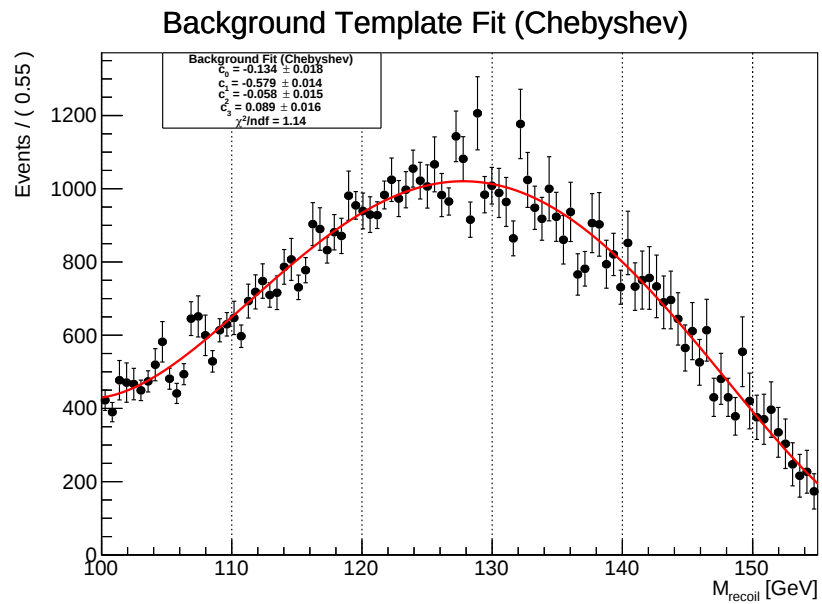


Рисунок 3.5 – Взвешенные распределения массы отдачи M_{recoil} для процессов $qqH\text{inv}$ (чистый сигнал $H \rightarrow 4\nu$, красная линия) и $qqHX$ (инклюзивный распад Хиггса, жёлтая гистограмма) после применения всех отборов



(a) Шаблон сигнала f_S (ядерная оценка плотности) без нормировки



(b) Шаблон фона f_B (полином Чебышёва 4-й степени) без нормировки

Рисунок 3.6 – Шаблоны сигнала и фона без нормировки, использованные для аппроксимации распределения массы отдачи M_{recoil}

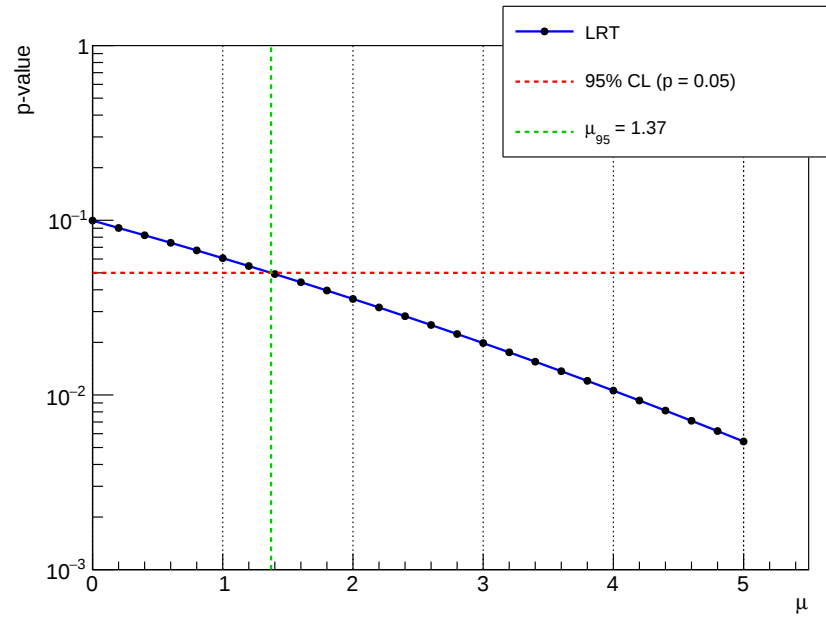


Рисунок 3.7 – Зависимость p -value для критерия отношения правдоподобий (LRT) от коэффициента усиления сигнала μ

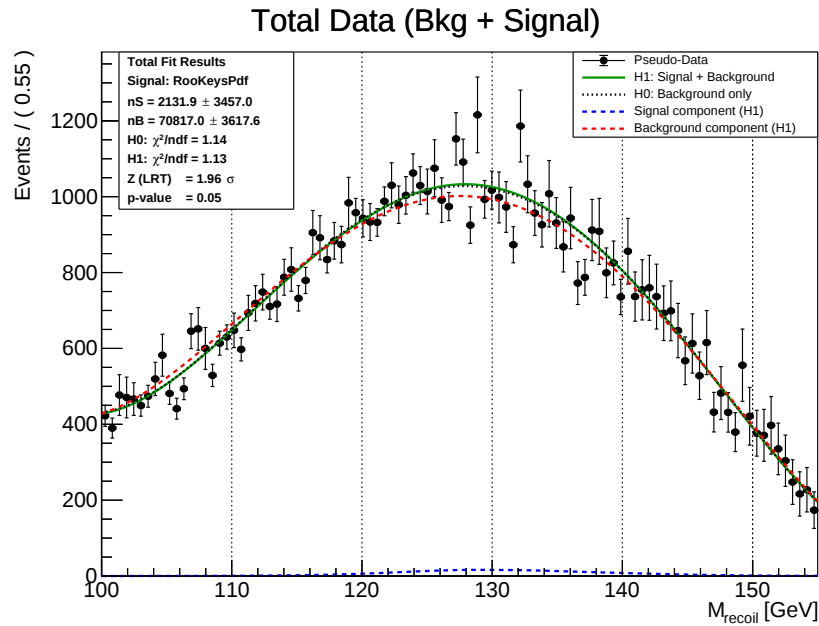
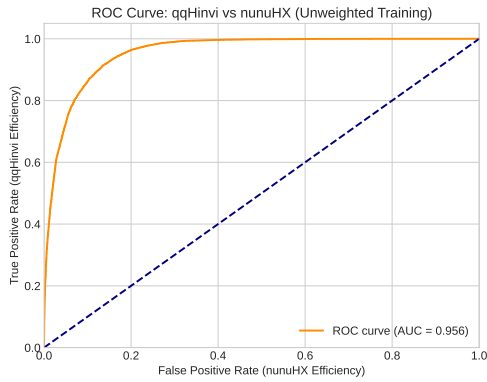
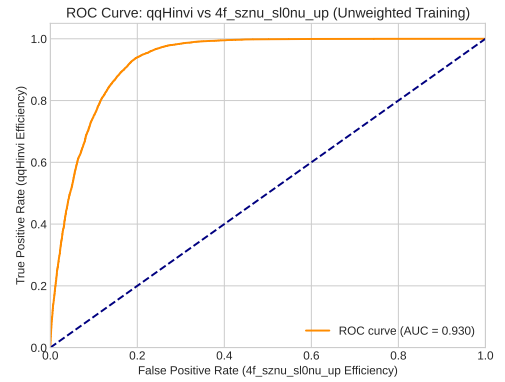


Рисунок 3.8 – Результат аппроксимации распределения массы отдачи M_{recoil} методом расширенного максимума правдоподобия. Псевдоданные (черные точки) сформированы из фоновых процессов и сигнального процесса $qq\text{Hinv}$, усиленного в $\mu = 1.37$ раз. Зелёная кривая: подгонка гипотезой H_1 (сигнал + фон); красный пунктир: компонента фона; синий пунктир: компонента сигнала. Черная пунктирная кривая: подгонка гипотезой H_0 (фон)

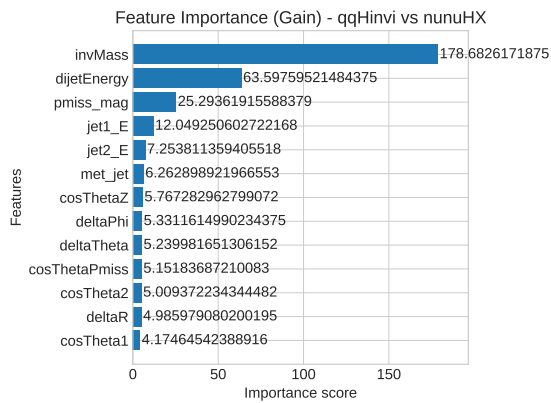


(a) qqHinvi vs nunuHX

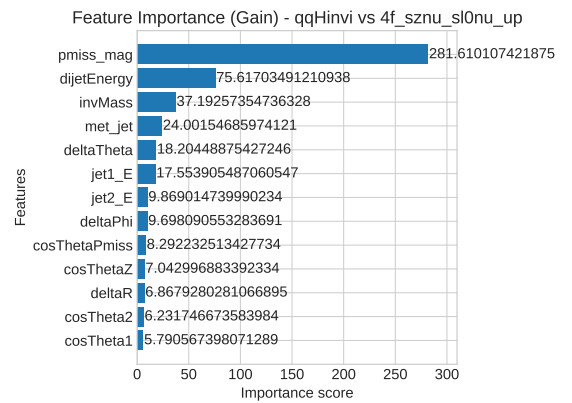


(b) qqHinvi vs 4f_sznusl0nu_up

Рисунок 3.9 – ROC-кривые обученных BDT-классификаторов для тестовой выборки

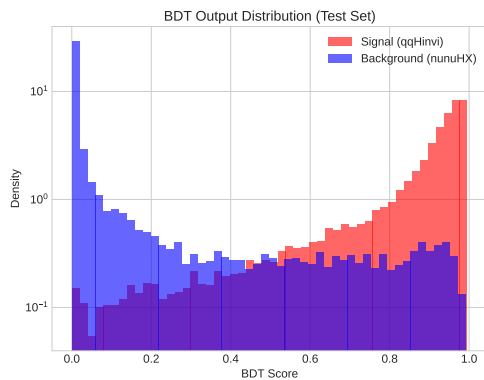


(a) qqHinvi vs nunuHX

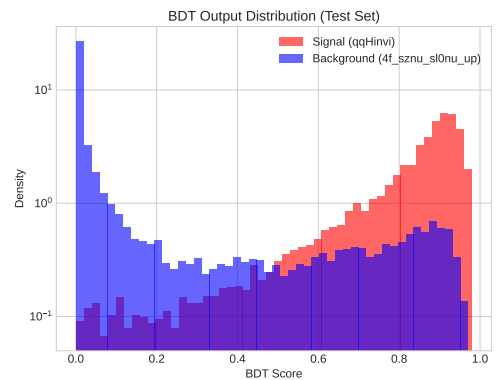


(b) qqHinvi vs 4f_sznusl0nu_up

Рисунок 3.10 – Важность признаков (по метрике gain) для обученных BDT-классификаторов



(a) qqHinvi vs nunuHX



(b) qqHinvi vs 4f_sznusl0nu_up

Рисунок 3.11 – Распределения выхода BDT-классификаторов для сигнальных и фоновых событий на тестовой выборке

ЗАКЛЮЧЕНИЕ

В данной работе проведён анализ потенциальной чувствительности эксперимента СЕРС к поиску распада бозона Хиггса на нерегистрируемые частицы в канале $e^+e^- \rightarrow ZH$ ($Z \rightarrow q\bar{q}$, $H \rightarrow \text{invisible}$). Основные результаты работы можно сформулировать следующим образом:

- изучена физическая мотивация поиска невидимых распадов Хиггса как канала для поиска тёмной материи и новой физики за пределами Стандартной Модели. Проведён обзор проекта СЕРС, его детекторной системы и программного обеспечения СЕРСВ;

- реализован алгоритм идентификации изолированных лептонов и фотонов, основанный на оценке энергии в конусе вокруг направления импульса частицы. Параметры алгоритма оптимизированы по результатам сканирования эффективности на контрольном процессе $e^+e^- \rightarrow ZH$ ($Z \rightarrow \mu^+\mu^-$, $H \rightarrow \text{inclusive}$), что позволило достичь эффективности $\varepsilon > 99\%$;

- применён обобщённый k_t -алгоритм кластеризации джетов (пакет FastJet) для реконструкции адронной системы от распада $Z \rightarrow q\bar{q}$;

- разработана двухэтапная стратегия отбора событий: предварительные отборы и основные кинематические отборы;

- выполнена аппроксимация распределения массы отдачи M_{recoil} методом расширенного максимума правдоподобия с использованием непараметрического шаблона на базе ядерной оценки плотности для сигнала и полинома Чебышева четвертой степени для фона. Найдено пороговое значение $\mu = 1.37$, которое соответствует $p\text{-value} = 0.05$ для критерия отношения правдоподобий. Это позволяет оценить верхний предел для измерения относительной вероятности распада бозона Хиггса на нерегистрируемые частицы на уровне 0.15% для доверительного интервала 95%;

- применены методы машинного обучения (BDT-классификатор) для повышения эффективности подавления фоновых процессов.

Необходимо отметить, что в текущем анализе учтены только статистические ошибки. Для получения физически интерпретируемых результатов необходимо исследовать основные источники систематической

погрешности.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC / G. Aad [и др.] // Physics Letters B. — 2012. — Т. 716, № 1. — С. 1—29. — ISSN 0370-2693. — DOI: 10.1016/j.physletb.2012.08.020. — URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
2. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC / S. Chatrchyan [и др.] // Physics Letters B. — 2012. — Т. 716, № 1. — С. 30—61. — ISSN 0370-2693. — DOI: 10.1016/j.physletb.2012.08.021. — URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
3. Implications of LHC searches for Higgs-portal dark matter / A. Djouadi [и др.] // Physics Letters B. — 2012. — Март. — Т. 709, № 1/2. — С. 65—69. — ISSN 0370-2693. — DOI: 10.1016/j.physletb.2012.01.062. — URL: <http://dx.doi.org/10.1016/j.physletb.2012.01.062>.
4. Combination of searches for invisible decays of the Higgs boson using 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment / G. Aad [и др.] // Physics Letters B. — 2023. — Т. 842. — С. 137963. — ISSN 0370-2693. — DOI: 10.1016/j.physletb.2023.137963. — URL: <http://dx.doi.org/10.1016/j.physletb.2023.137963>.
5. A search for decays of the Higgs boson to invisible particles in events with a top-antitop quark pair or a vector boson in proton-proton collisions at $\sqrt{s} = 13$ TeV / A. Tumasyan [и др.] // The European Physical Journal C. — 2023. — Окт. — Т. 83, № 10. — ISSN 1434-6052. — DOI: 10.1140/epjc/s10052-023-11952-7. — URL: <http://dx.doi.org/10.1140/epjc/s10052-023-11952-7>.
6. Cross section and Higgs mass measurement with Higgsstrahlung at the CEPC / Z.-X. Chen [и др.] // Chinese Physics C. — 2017. — Февр. — Т. 41, № 2. — С. 023003. — ISSN 1674-1137. — DOI: 10.1088/1674-1137/41/2/023003. — URL: <http://dx.doi.org/10.1088/1674-1137/41/2/023003>.

7. *Particle Data Group*. Review of Particle Physics: Higgs Boson. — 2023. — URL: <https://pdg.lbl.gov/2023/reviews/rpp2023-rev-higgs-boson.pdf> (дата обращения: 14.3.2026).
8. *Institute of High Energy Physics, Chinese Academy of Sciences*. CEPC. — 2025. — URL: <http://cepc.ihep.ac.cn/index.html> (дата обращения: 13.6.2025).
9. *Gao J.* CEPC Technical Design Report: Accelerator // Radiation Detection Technology and Methods. — 2024. — Март. — Т. 8, № 1. — С. 1—1105. — ISSN 2509-9949. — DOI: 10.1007/s41605-024-00463-y. — URL: <http://dx.doi.org/10.1007/s41605-024-00463-y>.
10. *The CEPC Study Group*. CEPC Conceptual Design Report: Volume 1 – Accelerator // arXiv e-prints. — 2018. — arXiv: 1809.00285 [physics.acc-ph].
11. *The CEPC Study Group*. CEPC Conceptual Design Report: Volume 2 – Physics and Detector // arXiv e-prints. — 2018. — arXiv: 1811.10545 [hep-ex].
12. *The CEPC Study Group*. CEPC Technical Design Report – Reference Detector // arXiv e-prints. — 2025. — arXiv: 2510.05260 [hep-ex].
13. *Thomson M.* Particle flow calorimetry and the PandoraPFA algorithm // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. — 2009. — Нояб. — Т. 611, № 1. — С. 25—40. — ISSN 0168-9002. — DOI: 10.1016/j.nima.2009.09.009. — URL: <http://dx.doi.org/10.1016/j.nima.2009.09.009>.
14. *Ruan M.* Arbor, a new approach of the Particle Flow Algorithm // arXiv e-prints. — 2014. — arXiv: 1403.4784 [physics.ins-det].
15. *CEPC Software Working Group*. CEPSCW Documentation. — 2026. — URL: <http://cepcsoft.ihep.ac.cn> (дата обращения: 14.4.2026).
16. GAUDI — A software architecture and framework for building HEP data processing applications / G. Barrand [и др.] // Computer Physics Communications. — 2001. — Окт. — Т. 140, № 1/2. — С. 45—55. — DOI: 10.1016/S0010-4655(01)00254-5.
17. *Ganis G., Helsens C., Völkl V.* Key4hep, a framework for future HEP experiments and its use in FCC. — 2021. — arXiv: 2111.09874 [hep-ex]. — URL: <https://arxiv.org/abs/2111.09874>.

18. DD4hep: A Detector Description Toolkit for High Energy Physics Experiments / M. Frank [и др.] // Journal of Physics: Conference Series. — 2014. — Т. 513, № 2. — С. 022010. — DOI: 10.1088/1742-6596/513/2/022010.
19. EDM4hep — a common event data model for HEP experiments / T. Madlener [и др.] // PoS. — 2022. — Т. ICHEP2022. — С. 1237. — DOI: 10.22323/1.414.1237.
20. Kilian W., Ohl T., Reuter J. WHIZARD—simulating multi-particle processes at LHC and ILC // The European Physical Journal C. — 2011. — Т. 71, № 9. — ISSN 1434-6052. — DOI: 10.1140/epjc/s10052-011-1742-y. — URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1742-y>.
21. A comprehensive guide to the physics and usage of PYTHIA 8.3 / C. Bierlich [и др.]. — 2022. — arXiv: 2203.11601 [hep-ph]. — URL: <https://arxiv.org/abs/2203.11601>.
22. Cacciari M., Salam G. P., Soyez G. FastJet user manual: (for version 3.0.2) // The European Physical Journal C. — 2012. — Март. — Т. 72, № 3. — ISSN 1434-6052. — DOI: 10.1140/epjc/s10052-012-1896-2. — URL: <http://dx.doi.org/10.1140/epjc/s10052-012-1896-2>.
23. HTCondor Team. HTCondor: High-Throughput Computing Software. — 2026. — URL: <https://htcondor.org> (дата обращения: 14.4.2026).
24. Verkerke W., Kirkby D. The RooFit toolkit for data modeling. — 2003. — arXiv: physics/0306116 [physics.data-an]. — URL: <https://arxiv.org/abs/physics/0306116>.
25. Cranmer K. Kernel estimation in high-energy physics // Computer Physics Communications. — 2001. — Май. — Т. 136, № 3. — С. 198—207. — ISSN 0010-4655. — DOI: 10.1016/S0010-4655(00)00243-5. — URL: [http://dx.doi.org/10.1016/S0010-4655\(00\)00243-5](http://dx.doi.org/10.1016/S0010-4655(00)00243-5).
26. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — ACM, 08.2016. — С. 785—794. — (KDD '16). — DOI: 10.1145/2939672.2939785. — URL: <http://dx.doi.org/10.1145/2939672.2939785>.