

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский ядерный университет «МИФИ»

УДК 53.05, 53.07

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

**Классификация лептонных распадов W бозона
методами машинного обучения в р-р столкновениях при
 $\sqrt{S} = 13$ TeV в эксперименте ATLAS**

Научный руководитель

_____ Д. Е. Пономаренко

Выполнил

_____ Г. А. Толкачёв

Москва 2021

Оглавление

1	Введение	2
1.1	Стандартная модель	3
2	Детектор ATLAS	4
3	Использованные данные	6
3.1	Экспериментальные данные	6
3.2	Монте-Карло моделирование	6
4	Предварительный отбор	8
4.1	Z регион	8
4.2	Сигнальный регион	9
5	Фитирование	14
5.1	Функция правдоподобия	14
5.2	Оптимизация биннирования BDT	17
5.3	Результат фитирования	18
6	Заключение	23
	Список использованных источников	24

1 Введение

Исследования в физике элементарных частиц привели к созданию теории взаимодействия частиц на субъядерном уровне, которую принято называть Стандартной моделью. В Стандартной модели существует три поколения лептонов. Согласно лептонной универсальности предполагается, что три поколения лептонов во всех процессах должны вести себя одинаково. В рамках данного предположения отношения отношений сечений лептонного распада W бозона для любой пары лептонных каналов распада должны быть равно единице. В анализе данных с LEP имеются различия между теоретическим предсказанием Стандартной модели и экспериментальными измерениями[6]. А именно, существует указание на возможное отклонение в отношении отношений сечения двух процессов лептонного распада W бозона ($Br(W \rightarrow \tau\nu) / Br(W \rightarrow \mu\nu)$) (рис. 1).

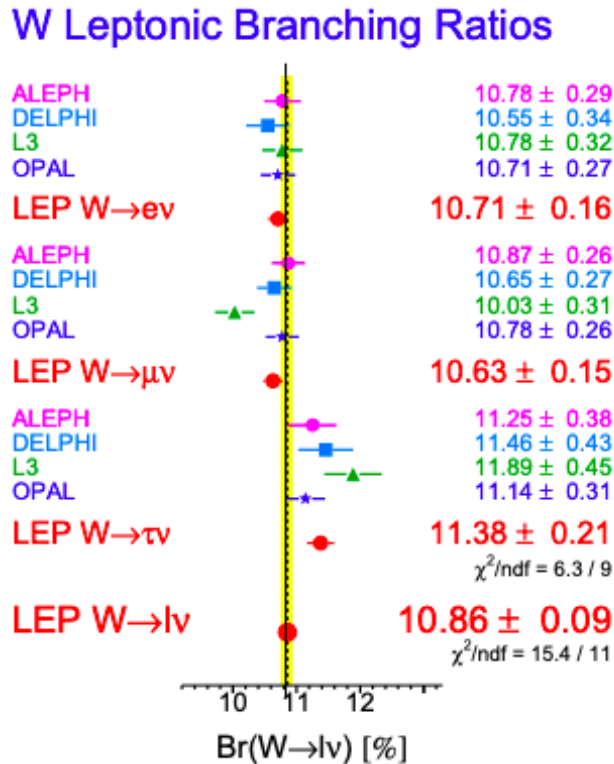


Рисунок 1 – Отношение сечения определенного канала распада к полному сечению распада W бозона

Анализ, в рамках которого проводится научно-исследовательская работа, является актуальным для поиска отклонения от предсказаний Стандартной

модели. Если в результате данного анализа полученный результат совпадет с теоретическим предсказанием, то это может свидетельствовать о подтверждении лептонной универсальности. В противном случае, если данный анализ подтвердит возможное отклонение в анализе данных с LEP, то это укажет на однозначное существование Новой физики. Данное, возможное отклонение от предсказаний Стандартной модели можно объяснить несколькими способами. Например, одно из возможных объяснений требует введения новой частицы, по-разному взаимодействующей с различными лептонами. Возможно, это новый класс частиц, например, лептокварки или же тяжелый аналог Z бозона.

Важным вкладом в данный анализ является работа над созданием и оптимизацией модели машинного обучения, с помощью которой можно провести классификацию данных. В результате классификации данных можно получить переменную, которая объединяет все кинематические переменные, используемые при классификации. На основе данной переменной удастся разделить очень кинематически схожие события. Таким образом, в дальнейшем можно будет провести более точное измерение отношения сечений лептонных распадов W бозона.

Целью данной работы является нахождение оптимальной модели машинного обучения лес деревьев решений (BDT). А также исследование влияния точности измерения лептонной универсальности от использования отклика классификатора BDT в базовом анализе.

1.1 Стандартная модель

Исследования в физике элементарных частиц привели к созданию теории взаимодействия частиц на субъядерном уровне, которую принято называть Стандартной моделью [13]. Стандартная модель позволяет теоретически предсказать свойства различных процессов в физике элементарных частиц. В рамках Стандартной модели имеется 2 типа элементарных частиц: бозоны и фермионы. Фермионы имеют полуцелый спин. Сами фермионы делятся на две подгруппы: кварки и лептоны. Лептоны делятся на 2 типа: электрически заряженные частицы (e , μ , τ) нейтральные частицы — нейтрино (ν_e ,

ν_μ, ν_τ). Кварки являются массивными частицами, имеющие электрический и цветовой заряд, всего их 6 (u, d, c, s, t, b). Все фермион по возрастанию массы делатся на три поколения. В стандартную модель входит 3 вида взаимодействия: электромагнитное, слабое и сильное. Сильное взаимодействие описывается квантовой хромодинамикой(КХД). Электромагнитное и слабое взаимодействия являются составными частями электро-слабого взаимодействия.

Лептонная универсальность

В рамках Стандартной модели существует три поколения лептонов. Три поколения упорядочены по массе заряженного лептона в диапазоне от 0.511 МэВ для e до 105 МэВ для μ , и 1.777 для τ [11]. Различие в массе приводят к совершенно разным временам жизни. От стабильного e до 2.2 мкс для μ , и 0.29 пс для τ . Лептоны участвуют в электромагнитных и слабых, но не сильных взаимодействиях, тогда как нейтрино участвуют только в слабом взаимодействии. Стандартная модель предполагает, что эти взаимодействия заряженных и нейтральных лептонов универсальны, т.е. одинаковы для трех поколений. Количество поколений лептонов пока не объяснено в рамках стандартной модели. Почти все наблюдаемые во Вселенной процессы выглядели бы точно так же, если бы существовало только одно поколение лептонов. Лептонную универсальность можно, например, проверить в лептонном распаде W бозона. Масса W достаточно большая, чтобы пренебречь массами лептонов, на которые он распадается. Поэтому между лептонными распадами W не должно быть разницы.

2 Детектор ATLAS

ATLAS (от англ. A Toroidal LHC ApparatuS) — один из четырёх основных экспериментов на коллайдере LHC в Европейской Организации Ядерных исследований CERN в городе Женева (Швейцария) (рис. 2). Эксперимент проводится на одноимённом детекторе, предназначенном для исследования протон-протонных столкновений. Детектор состоит из нескольких частей. Для восстановления треков и импульсов заряженных частиц используется внутренний детектор, окруженный сверхпроводящим магнитом, создающим

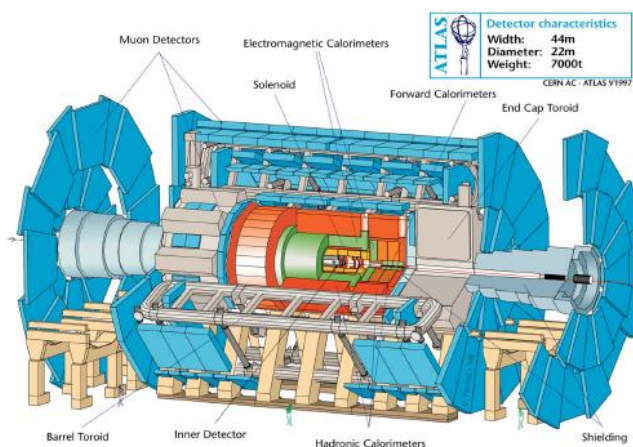


Рисунок 2 – Детектор ATLAS

магнитное поле 2 Тл. С помощью системы калориметров происходит измерение энергии частиц. На периферии детектора находится мюонный спектрометр, который необходим для измерения импульса мюонов. Для отбора событий используется система триггеров.

3 Используемые данные

Во время набора данных на эксперименте ATLAS рождается огромное число фоновых событий, для того чтобы понять какой вклад они вносят применяют статистический анализ, важным компонентом которого являются данные, сгенерированные методом Монте-Карло[9]. В работе использовались данные, полученные с использованием *HistMaker*[1].

3.1 Экспериментальные данные

В работе использовались экспериментальные данные, набранные на детекторе ATLAS в 2017 и 2018 году во время режима набора данных с низкой светимостью 340 пБ^{-1} . При столкновении протон-протонных пучков с суммарной энергией 13 ТэВ.

3.2 Монте-Карло моделирование

Смоделированные данные, используемые в работе, были получены методом Монте-Карло с помощью генераторов Pythia[12] и Sherpa[7] и прошли всю цепочку реконструкций, на условии реальных протон-протонных столкновений эксперимента ATLAS сессия 2 (RUN 2). Каждому каналу соответствует свой уникальный номер DSID. Список Монте-Карло данных, использованных в работе, приведен в таблице 1

Для сравнения Монте-Карло и реальных данных выполняется нормировка на светимость. Для более точного согласия с распределениями из данных используются коррекционные коэффициенты, которые учитывают неточности в моделирование Монте-Карло и геометрию детектора. Коэффициенты коррекции, использованные в работе предоставляются Combined Performance (CP) Groups эксперимента ATLAS [5].

Sample	DSID	Generator	xs [pb]
$W^+ \rightarrow \mu\nu$	361101	PowhegPythia8EvtGen	11500.9
$W^- \rightarrow \mu\nu$	361104	PowhegPythia8EvtGen	8579.31
$W^+ \rightarrow \tau\nu$	361102	PowhegPythia8EvtGen	11500.9
$W^- \rightarrow \tau\nu$	361105	PowhegPythia8EvtGen	8579.31
$Z \rightarrow \tau\tau$	361108	PowhegPythia8EvtGen	1950.63
$Z \rightarrow \mu\mu$	361107	PowhegPythia8EvtGen	1950.63
Top	410013	PhPy8EG_P2012	35.8455
Top	410014	PhPy8EG_P2012	35.8244
Top	410470	PhPy8EG	729.77
Top	410642	PhPy8EG	36.993
Top	410643	PhPy8EG	22.174
Top	410644	PowhegPythia8EvtGen	2.06146
Top	410645	PowhegPythia8EvtGen	1.28867
Diboson	363356	Sherpa_221_PDF30	2.20355
Diboson	363358	Sherpa_221_PDF30	3.4328
Diboson	363359	Sherpa_221_PDF30	24.708
Diboson	363360	Sherpa_221_PDF30	24.724
Diboson	363489	Sherpa_221_PDF30	11.42
Diboson	364250	Sherpa_221_PDF30	1.2523
Diboson	364253	Sherpa_221_PDF30	4.579
Diboson	364254	Sherpa_221_PDF30	12.501
Diboson	364255	Sherpa_221_PDF30	3.2344

Таблица 1 – Список каналов, использованных в генераторе Монте-Карло при моделировании данных.

4 Предварительный отбор

В данной работе используется два контрольных региона. Один является сигнальным регионом, а второй является Z регионом. Каждый регион характеризуется отдельным набором ограничений на кинематические переменные и сигнатуру событий.

4.1 Z регион

В данном анализе Z регион используется как контрольный регион. В нем экспериментальные данные очень хорошо согласуются с Монте-Карло данными, а также содержится малое количество КХД фона. С помощью Z региона проводят валидацию сигнального региона и используют для уменьшения систематической погрешности.

Отбор событий Z региона

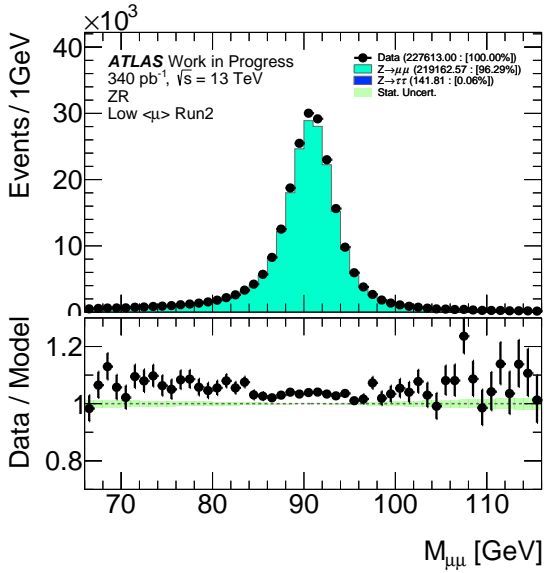
Для получения Z региона был проведен отбор, который включает в себя ограничение на поперечный импульс P_T больше 20 ГэВ для двух лептонов. Отбор на инвариантную массу от 66 до 116 ГэВ. А также были применены критерии на изоляцию лидирующего лептона и идентификацию.

P_T	M_{ll}	ID	Isolated
> 20 ГэВ	66 ГэВ $< m < 116$ ГэВ	tight	medium

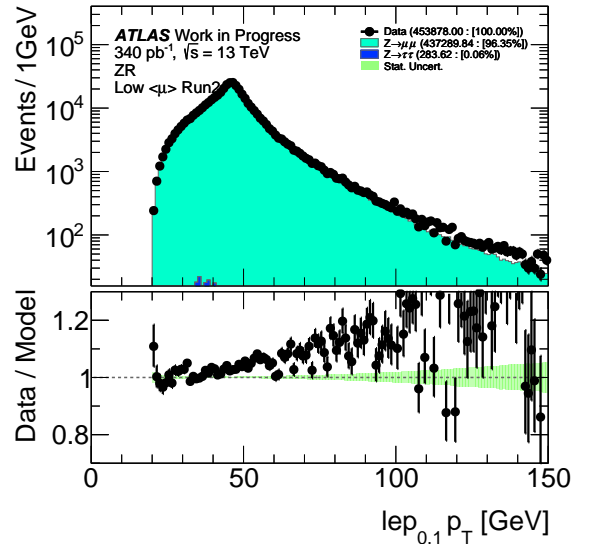
Таблица 2 – Отбор событий Z региона

Кинематические распределения Z региона

Для Z региона было выполнено сравнение кинематических распределений для реальный и Монте—Карло данных, а также показано их отношение. В Z регионе содержится мало КХД фона и Монте—Карло хорошо согласуются с экспериментальными данными, однако, на представленном отношении



(а) Инвариантная масса $M_{\mu\mu}$



(б) Поперечный импульс P_T каждого лептона из пары

Рисунок 3 – Сравнение реальных и Монте-Карло данных Z региона для переменных: $M_{\mu\mu}$ и поперечного импульса каждого лептона P_T из пары

имеются отклонения Монте—Карло от экспериментальных данных в жесткой части поперечного импульса каждого лептона из пары $lep_{0,1}P_T$ (рис. 3б).

4.2 Сигнальный регион

Сигнальный регион является основным в данной работе. Он содержит события с распадами W бозона на τ лептон. На данных сигнального региона производится обучение модели машинного обучения. А также с использованием данных сигнального региона получают отклик модели, который в дальнейшем планируется использовать для финального измерения лептонной универсальности в базовом анализе. Для отбора сигнального региона отбираются события, которые имеют схожую сигнатуру с распадом W бозона в τ лептон.

Отбор событий сигнального региона

Для получения сигнального региона был проведен отбор, который включает в себя ограничение на поперечный импульс P_T больше 20 ГэВ. Отбор на поперечную энергию E_T^{miss} больше 20 ГэВ, поперечную массу M_T больше 40

ГэВ, определяемую как

$$m_T = \sqrt{2P_T(l)P_T(\nu)(1 - \cos(\phi(l) - \phi(\nu)))}, \quad (1)$$

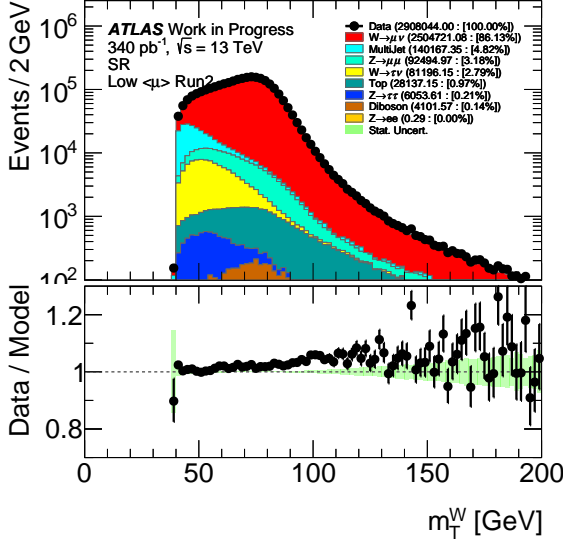
Данный отбор проводится для отсеечения большого количества КХД фона. Также использовалось ограничение на количество лептонов в событии равное 1. Были применены критерии идентификацию мюона - Tight, а также на изолированность мюона - Medium(таблица 3). Для улучшения выделения настоящих лептонов от других, например, не интересующих нас частиц или струй, к лептонам применяется отбор на изолированность. Трековая изоляция рассчитывается как сумма поперечных импульсов всех треков в конусе размера $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, кроме трека самого лептона, и делится на импульс лептона. В работе используется трековая изоляция в конусе $\Delta R = 0.2$, равная $ptvarcone20/P_T < 0.1$.

P_T	E_T^{miss}	m_T	N lep	$ptvarcone20/P_T$	ID	Isolated
> 20 ГэВ	> 20 ГэВ	> 40 ГэВ	1	< 0.1	tight	medium

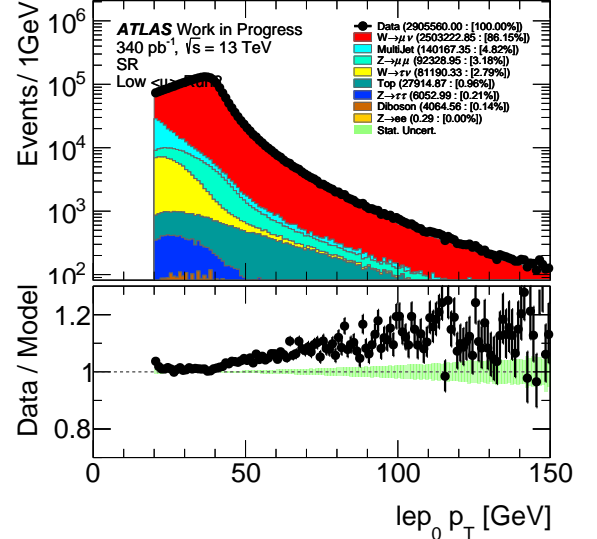
Таблица 3 – Отбор событий сигнального региона

Кинематические распределения сигнального региона

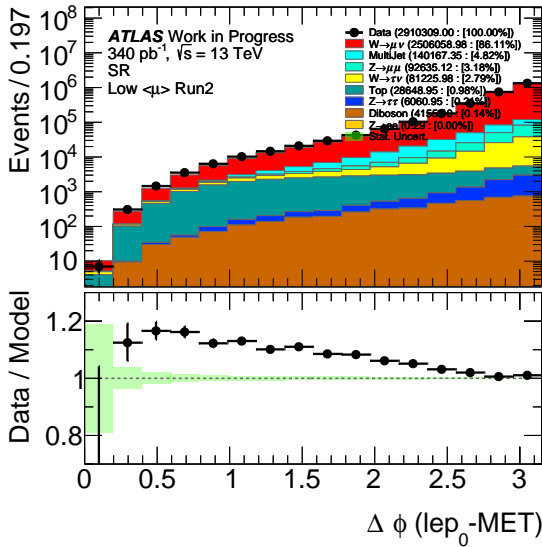
Было выполнено сравнение кинематических распределений для реальных и Монте-Карло данных (рис. 4 и 5), а также показано отношение реальных данных к данным Монте-Карло. По данному распределению видно, что по-



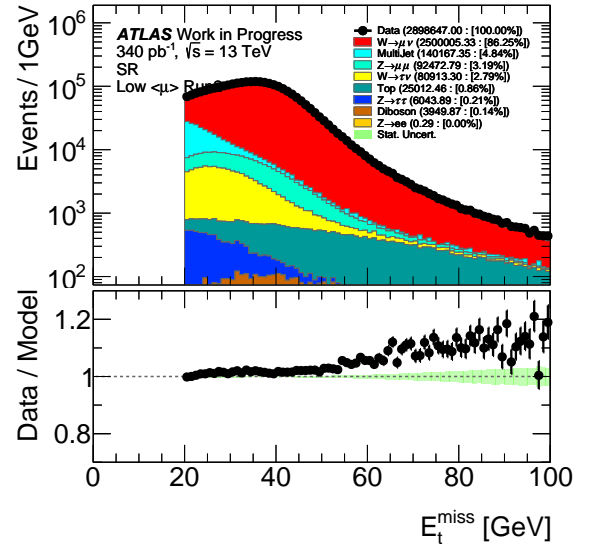
(а) Поперечная масса m_T



(б) Поперечный импульс P_T



(в) Разность углов ϕ потерянной энергии и лептона $d\phi(l - E_{\text{miss}})$

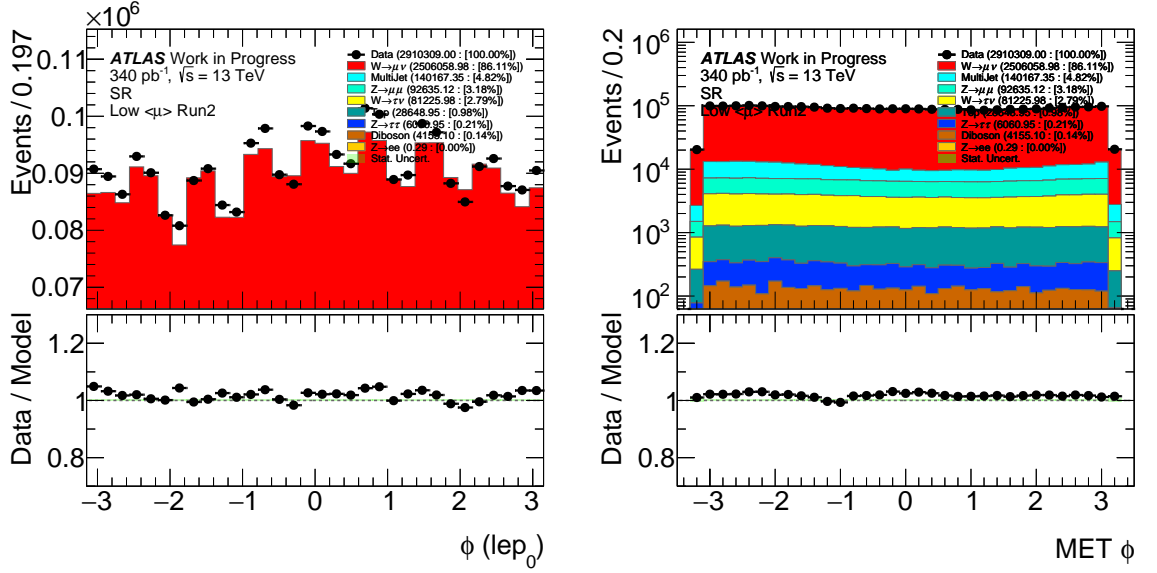


(г) Потерянная поперечная энергия E_{miss}

Рисунок 4 – Сравнение реальных и Монте-Карло данных для переменных: M_T , P_T , $d\phi(l - E_{\text{miss}})$, E_{miss}

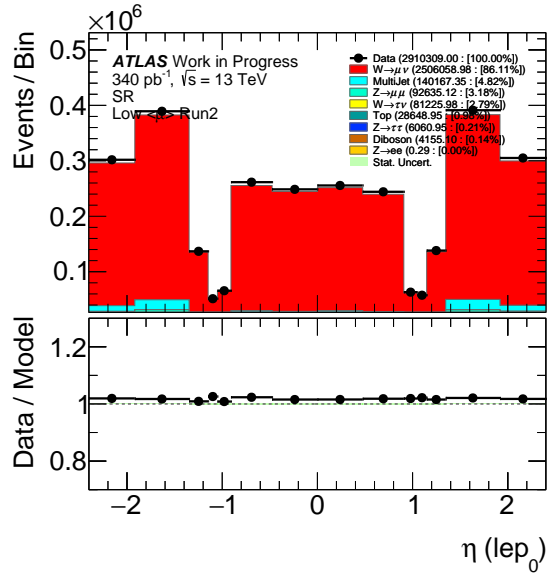
мимо сигнального процесса большой вклад вносят фоновые события. Наибольший вклад вносит процесс распада W бозона в лептон второго положе-

ния, значительный вклад вносит также КХД фон и распады Z бозона на два лептона. Участие распада Z бозона в фоновом процессе происходит из-за того, что иногда один из лептонов не удается зарегистрировать. Поэтому данный распад имеет похожую сигнатуру с распадом W бозона в лептон. На представленном отношении отчетливо видно отклонение смоделирован-



(а) Угол ϕ лептона

(б) Угол ϕ потерянной энергии



(в) Псевдобыстрота η

Рисунок 5 – Сравнение реальных и Монте-Карло данных для переменных: $\text{lep } \phi$, MET ϕ , $\text{lep } \eta$

ных Монте-Карло данных от реальных данных в жесткой части поперечного импульса P_T , потерянной энергии E_T^{miss} , поперечной массы M_T . Данное отклонение имеется также в Z регионе. Поэтому можно сделать вывод о том,

что данное отклонение связано с плохим моделированием Монте-Карло в генераторе PowhegPythia.

Оценка КХД фона

Используемые в данной работе генераторы Монте—Карло не могут надежно смоделировать КХД фон. Поэтому необходимо произвести оценку КХД фона с помощью метода оценки фона из данных(data-driven method)[10]. КХД фон был посчитан анализ группой[2] и предоставлен в мою работу для использования в исследовании отклика VDT моделей.

5 Фитирование

В работе проводится измерение отношения отношений сечений с помощью биннированного метода максимального правдоподобия. Сравниваются результаты фитирования с использованием двух разных переменных, а именно поперечной массы M_T и переменной отклика модели BDT, которая была получена в ходе работы предыдущего семестра.

Первоначально фитирование необходимо провести в статусе слепого анализа, поэтому вместо реальных данных используются данные, которые полностью соответствуют Монте-Карло симуляции, при этом имеют статистическую погрешность реальных данных. Такие данные называются Азимов данными. Исходя из того, что используются данные, которые полностью соответствуют Монте-Карло симуляции можно сделать вывод о том, что полученное значение отношения отношений сечений после фитирования будет равняться единице. На данном этапе работы интересным является не само значение отношения отношений сечений, а его статистическая и систематическая погрешность, которую предполагается снизить за счет использования переменной отклика модели BDT.

Предполагается, что с помощью переменной отклика модели BDT удастся более точно разделить сигнальные и фоновые события. Эти предположения были выдвинуты за счет того, что переменная отклика модели BDT комбинирует в себе набор кинематических переменных, в том числе поперечную массу M_T , с помощью которой можно достаточно хорошо разделить сигнальные и фоновые события.

5.1 Функция правдоподобия

Для построения функции правдоподобия и дальнейшего проведения фитирования производится создание гистограмм с помощью xTauReader [3]. Получение параметра интереса производится с помощью TRexFitter [4]. Перед началом проведения фитирования необходимо правильно задать функцию правдоподобия и параметр интереса. Параметром интереса в работе является отношение отношений сечений $R_{\tau/\mu}$.

$$R_{\tau/\mu} = \frac{BR(W \rightarrow \tau\nu)}{BR(W \rightarrow \mu\nu)} \quad (2)$$

Покажем, что параметр интереса является отношением отношений сечений. Экспериментальные данные сигнального региона можно выразить через Монте-Карло данные соответствующей формулой,

$$N_{data} = \mu_W(\mu_{sig}N_\tau + N_l) + \mu_{ZR}N_{MC} + N_{QCD} \quad (3)$$

где переменная μ_{sig} является параметром интереса, а μ_W и μ_{ZR} нормировочными коэффициентами. Вклад, который вносит распады $W \rightarrow \tau\nu$ и $W \rightarrow \mu\nu$ обозначается как N_τ и N_l , а за вклад КХД фона отвечает N_{QCD} . Доля всех остальных процессов может объединена и обозначена как N_{MC} . Использование нормировочных коэффициентов μ_W и μ_{ZR} оправдано тем, что возможны отклонения Монте-Карло симуляций от реальных данных, поэтому для учета данного отклонения необходимо задать нормировочные коэффициенты. В противном случае, без учета этих нормировочных параметров, параметр интереса будет оценен неверно.

Обозначим отношение отношений сечений распада $W \rightarrow \tau\nu \rightarrow \mu\nu\nu$ к $W \rightarrow \mu\nu$ за $R_{\tau/\mu}^{exp}$,

$$R_{\tau/\mu}^{exp} = \frac{BR(W \rightarrow \tau\nu)BR(\tau \rightarrow \mu\nu\nu)}{BR(W \rightarrow \tau\nu)} \quad (4)$$

Тогда, параметр интереса можно задать соответствующим образом.

$$\mu_{sig} = \frac{\left(\frac{BR(W \rightarrow \tau\nu)BR(\tau \rightarrow \mu\nu\nu)}{BR(W \rightarrow \tau\nu)}\right)_{data}}{\left(\frac{BR(W \rightarrow \tau\nu)BR(\tau \rightarrow \mu\nu\nu)}{BR(W \rightarrow \tau\nu)}\right)_{MC}} \quad (5)$$

Некоторые параметры при генерировании данных Монте-Карло могут быть заданы исходя из результатов предыдущих экспериментов. Например, в работе для данных Монте-Карло было установлено $BR(\tau \rightarrow \mu\nu\nu)$ равное 17.39 процентам. А с учетом того, что для экспериментальных данных $BR(\tau \rightarrow \mu\nu\nu)$ соответствует тому же значению, то на эту величину можно сократить числитель и знаменатель. Таким образом получим,

$$\mu_{sig} = \left(\frac{BR(W \rightarrow \tau\nu)}{BR(W \rightarrow \mu\nu)}\right)_{data} / \left(\frac{BR(W \rightarrow \tau\nu)}{BR(W \rightarrow \mu\nu)}\right)_{MC} \quad (6)$$

Так как симуляция Монте-Карло основывается на теоретических предсказаниях стандартной модели можно приравнять отношение $\left(\frac{BR(W \rightarrow \tau\nu)}{BR(W \rightarrow \mu\nu)}\right)_{MC}$ к единице.

$$\mu_{sig} = R_{\tau/\mu} = \frac{BR(W \rightarrow \tau\nu)}{BR(W \rightarrow \mu\nu)} \quad (7)$$

Из чего можно заключить, что параметр интереса μ_{sig} равен отношению отношений сечений $R_{\tau/\mu}$.

Для фитирования используется глобальная функция правдоподобия, которая включает в себя бины и же все фитируемые параметры μ_{sig} , μ_W и μ_{ZR} . Фитирование настраивается таким образом, чтобы максимизация функции правдоподобия выполнялась с несколькими неизвестными параметрами, которые представляют интерес для данного анализа, а также с систематическими погрешностями, которые включены в качестве мешающих параметров θ . Мешающие параметры являются параметрами, которые не представляют непосредственного интереса, но которые должны быть учтены при анализе параметров, представляющих интерес. Сама функция правдоподобия определяется стандартным способом в виде произведения распределения Пуассона для каждого бина, каждого используемого региона и функции плотности вероятности G , которая учитывает систематические погрешности.

$$L = \prod_{bins \in ZR} Pois(N_b^{data}, \mu_{ZR} N_b^{MC}) \cdot \prod_{bins \in SR} Pois(N_b^{data}, \mu_W [\mu_{sig} N_b^T + N_b^I] + \mu_{ZR} N_b^{MC} + N_b^{QCD}) \prod_j^{NPs} G(\theta_j^0, \theta_j) \quad (8)$$

Фитирование задается с тремя неопределенными параметрами: μ_{sig} - интересующий нас параметр, и нормировочные коэффициенты μ_W и μ_{ZR} . Нормировочные коэффициент для КХД фона не является неопределенным из за того, что при получении КХД фона был использован метод оценки КХД фона из экспериментальных данных. Однако, стоит учесть систематические погрешности КХД фона в виде мешающих параметров. Для КХД фона имеется две систематические погрешности, которые входят в фит как мешающие

параметры. Первой является систематическая погрешность, которая меняется от бина к бину. Второй является погрешность на форму распределения КХД фона, которая не является зависимой от бина к бину. Все остальные систематические погрешности также как и погрешности для КХД фона входят в функцию правдоподобия как мешающие параметры. В работе в функцию G входят только мешающие параметры, отвечающие за статистическую погрешность Монте-Карло симуляций. Мешающие параметры, отвечающие за систематические погрешности оценки КХД фона планируется включить в работе позже.

5.2 Оптимизация биннирования BDT

Перед началом проведения фитирования была произведена оптимизация биннинга для переменной отклика модели BDT. Оптимизация биннинга проводится таким образом, чтобы минимизировать систематическую и статистическую погрешность для параметра интереса μ^{sig} . Двухмерный фит переменных отклика BDT и прицельного параметра d_0 выполняется путем разбиения событий на 5 областей для отклика BDT, каждая из которых имеет соответствующее одномерное распределение d_0 , которое в свою очередь разбито на 8 бинов. Также производится разбиение инвариантной массы Z региона на 50 бинов. Таким образом общее количество бинов, которые участвуют в фитировании составляет $8 \cdot 5 + 50 = 90$.

Разбиение распределения BDT на области выбрано таким образом, чтобы получить максимальную чувствительность к изменению формы сигнального и фонового распределения. При выборе разбиения учитывается также отношение количества сигнальных событий к фоновым. На рис. 6 продемонстрировано разбиение переменной отклика классификатора BDT с учетом формы распределения и отношения сигнальных событий к фоновым. Синими линиями обозначены разбиения распределения отклика классификатора BDT на бины.

Оптимизация биннирования приводит к следующему разбиению для отклика BDT [0.0, 0.3, 0.5, 0.65, 0.83, 1.0].

Разбиение для d_0 было взято из аналогичного анализа $t\bar{t}$ группы [8]. Для

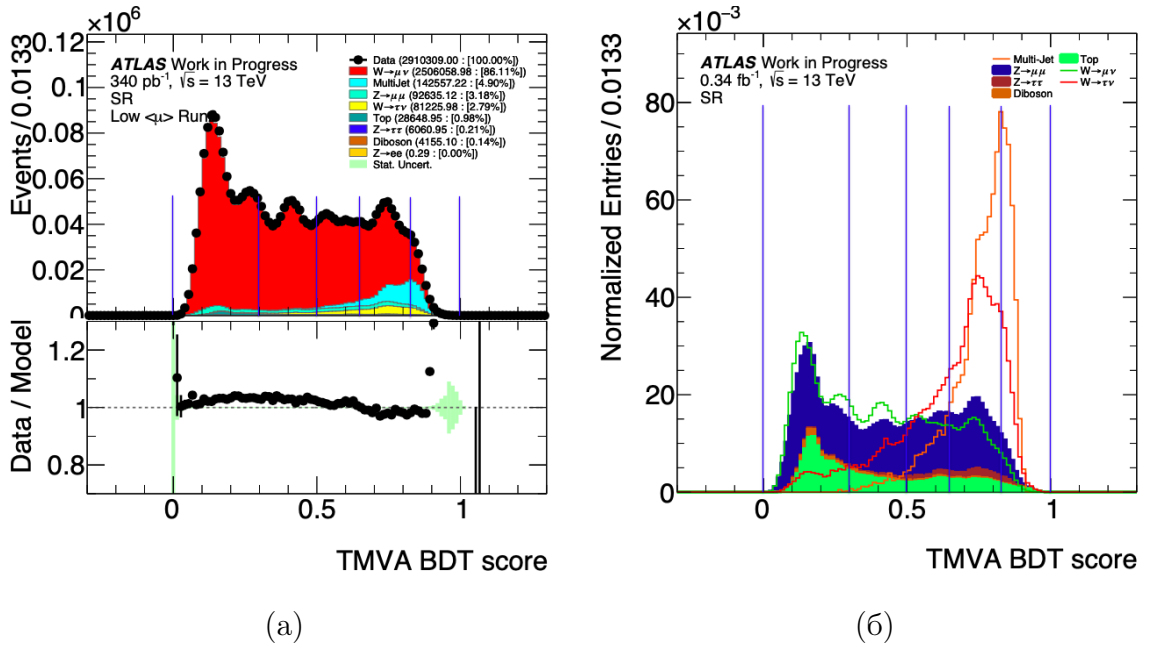


Рисунок 6 – (а) - Распределение отклика классификатора BDT, (б) - Нормированное распределение отклика классификатора BDT

d_0 так же как и для отклика BDT разбиение выбрано таким образом, чтобы выделить сигнальные события $[0, 0.01, 0.02, 0.03, 0.04, 0.06, 0.09, 0.15, 0.5]$. В работе не проводилась оптимизация бининга для d_0 . Это связано с тем, что исследуется зависимость систематической и статистической погрешности отношения отношений сечений от использованной переменной при постоянном разбиении для d_0 .

5.3 Результат фитирования

В работе проводится исследование влияния используемой переменной на статистическую погрешность во время фитирования. Исследуются две переменные: поперечная масса M_T и отклик модели BDT. Предполагается, что с использованием BDT отклика в фитировании обеспечит более точное измерение отношения отношений сечений, так как BDT формируется на основе 7 переменных, одной из которых является поперечная масса M_T . Фитирование проводится на Азимов данных, которые используются для проверки работоспособности фита, а также для выявления различных проблем, таких как сильно ограниченные мешающие параметры или большие корреляции между параметрами фита. Азимов данные также используются для определения то-

го, какие погрешности оказывают наибольшее влияние на параметр интереса μ_{sig} .

На первом этапе работы было произведено фитирование с использованием поперечной массы. Разбиение переменной поперечной массы M_T не было оптимизировано в работе, а было взято из базового анализа. На рисунке 7а продемонстрировано разбиение переменной поперечной массы M_T на бины, а на рисунке 8 показано распределение переменной $d0$ в каждом бине M_T .

После проведения фитирования было получено значение параметра интереса μ^{sig} и его статистической погрешности.

$$\mu_{sig} = 1.0000_{-0.0121}^{+0.0121}(stat)_{-0.0051}^{+0.0051}(stat_{MC})$$

На втором этапе по аналогии с первым было проведено фитирование с использованием новой переменной отклика BDT. На рисунке 7б показано разбиение для переменной отклика модели BDT, а на рисунке 9 показано распределение переменной $d0$ в каждом бине BDT.

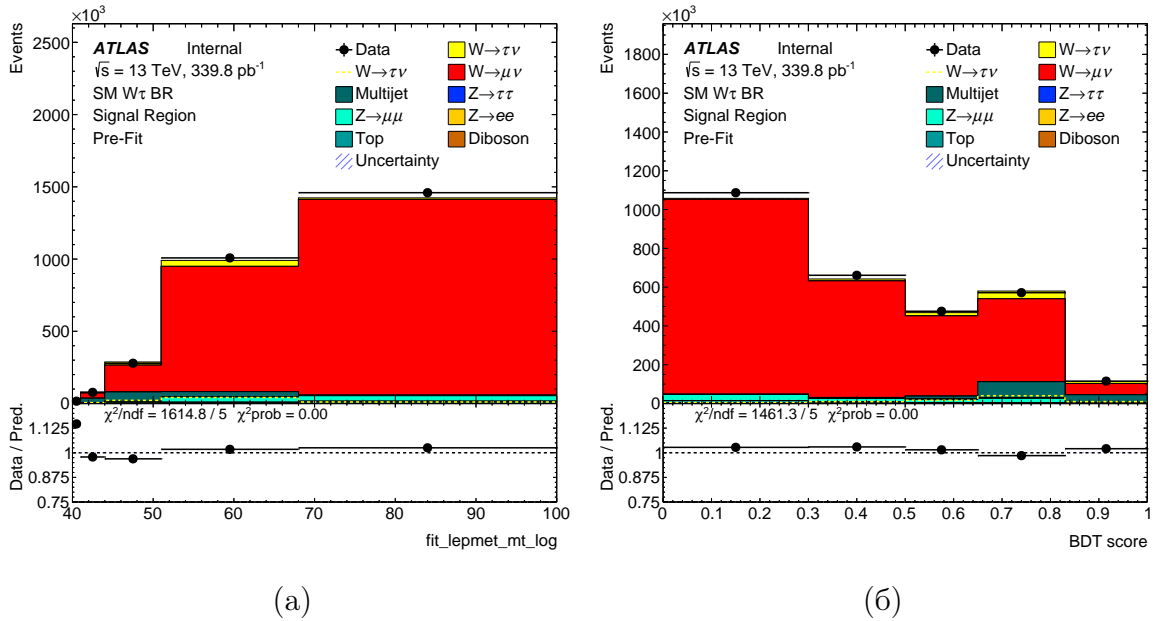
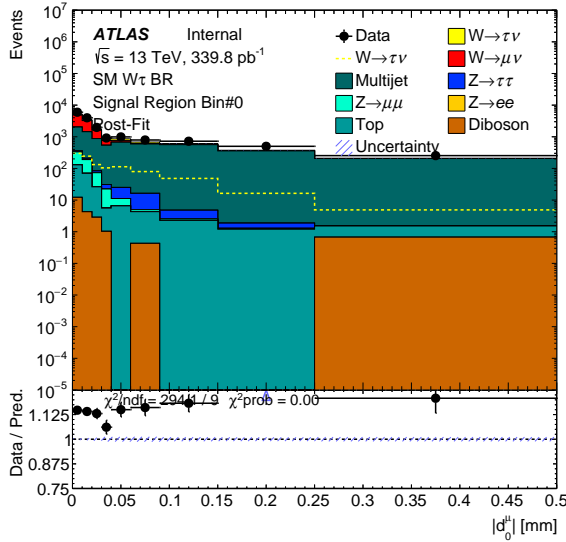


Рисунок 7 – (а) - Распределение переменной поперечной массы M_T , (б) - распределение переменной отклика модели BDT

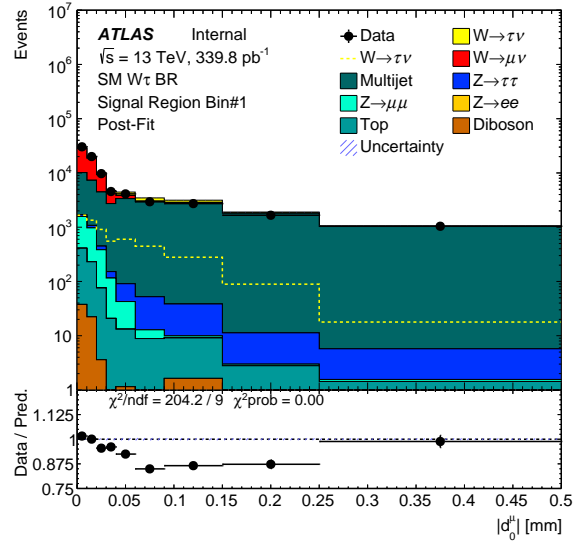
Получено значение параметра интереса и его статистической погрешности.

$$\mu_{sig} = 1.0000_{-0.0117}^{+0.0117}(stat)_{-0.0048}^{+0.0048}(stat_{MC})$$

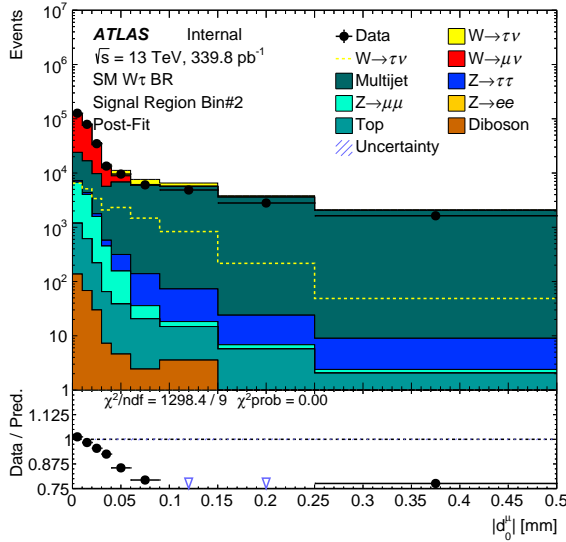
Сравнение двух результатов фитирования говорит о том, что использование переменной отклика модели при фитировании уменьшает статистическую погрешность данных на 0.004 и статистическую погрешность симуляций Монте-Карло погрешность на 0.003. Однако, различие между двумя результатами фитирования является несущественным. В дальнейшем планируется исследование причины столь малого уменьшения погрешностей, а также проведение повторного фитирования и сравнения результатов.



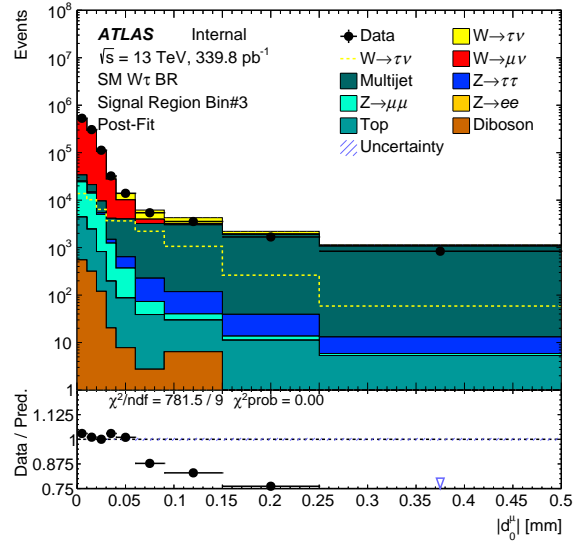
(a)



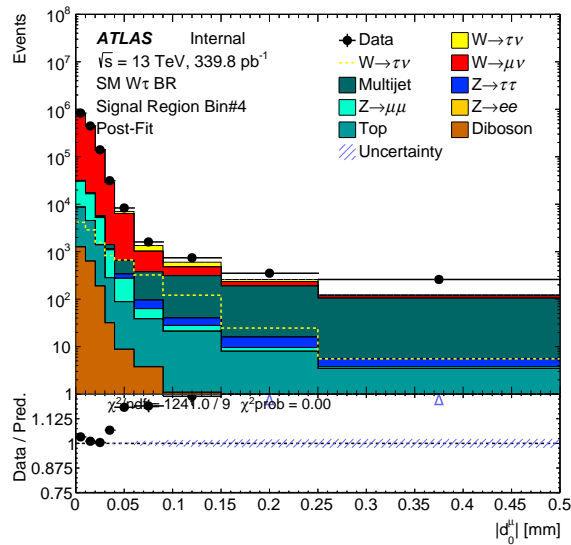
(б)



(B)

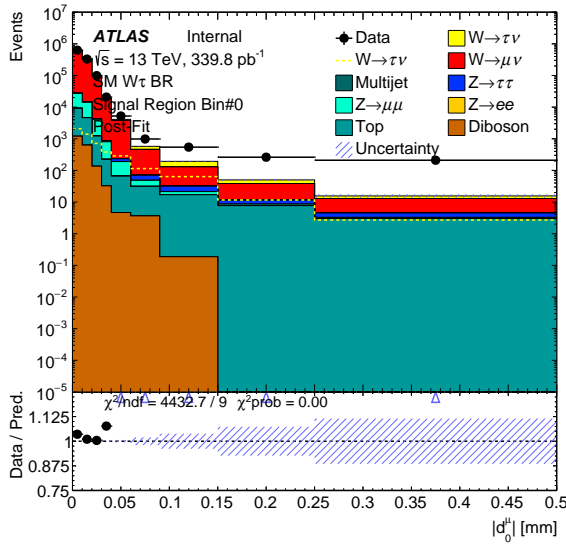


(Г)

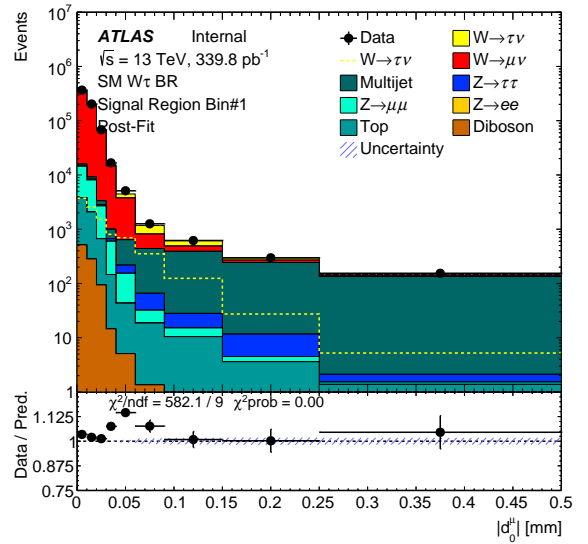


(д)

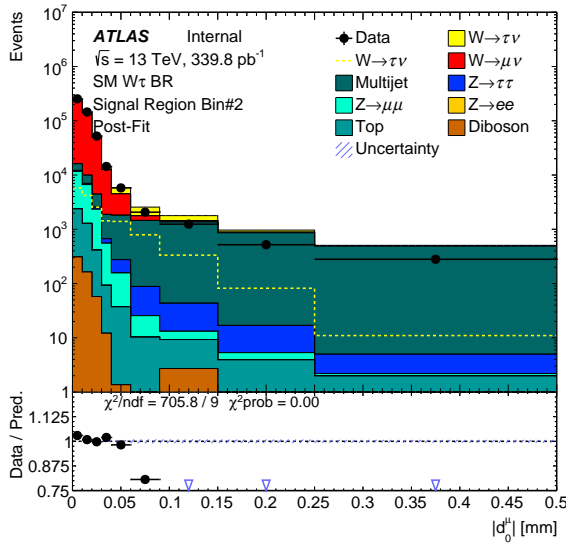
Рисунок 8 – Разбиение переменной d_0 по бинам переменной поперечной массы M_T



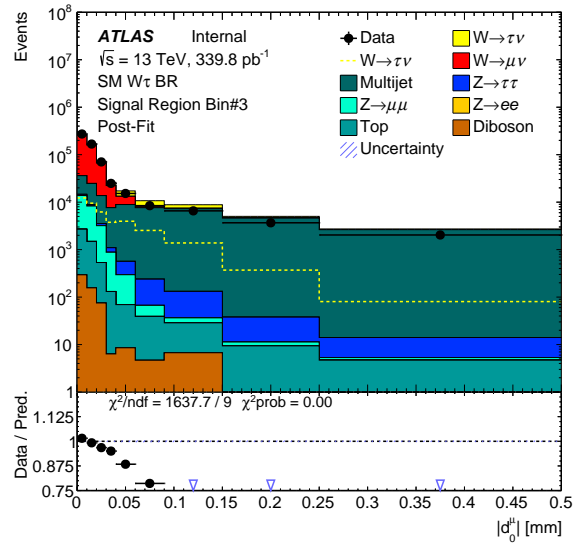
(a)



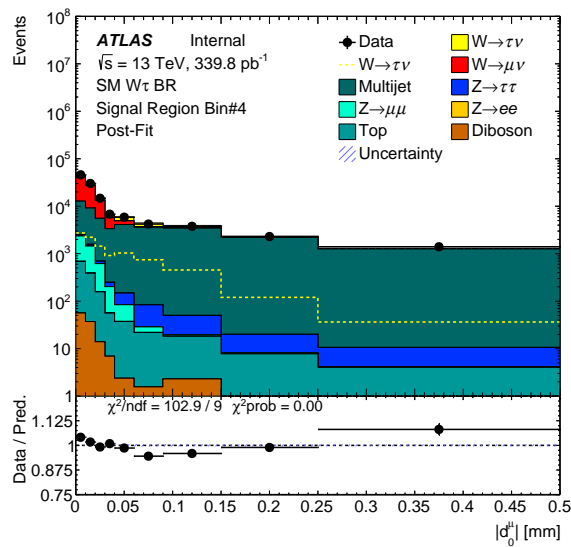
(б)



(B)



(Г)



(д)

Рисунок 9 – Разбиение переменной d_0 по бинам переменной BDT

6 Заключение

В процессе проведения работы произведен вывод функции правдоподобия и параметра интереса, которым является отношение отношений сечений. Для уменьшения погрешностей параметра интереса и для максимального разделения фоновых и сигнальных событий проведена оптимизация биннирования новой переменной отклика модели BDT. Таким образом был произведен выбор оптимального биннинга для переменной отклика классификатора BDT [0.0, 0.3, 0.5, 0.65, 0.83, 1.0].

Выполнено фитирование с использованием с новой переменной отклика классификатора BDT. Получены значение параметра интереса и его погрешность, в которую входит статистическая погрешность экспериментальных и Монте-Карло данных. В случае использования новой переменной отклика классификатора BDT .

$$\mu_{sig} = 1.0000^{+0.0117}_{-0.0117}(stat_{data})^{+0.0048}_{-0.0048}(stat_{MC})$$

Результат, полученный с использованием новой переменной отклика модели BDT не сильно отличается от результата, который был получен с использованием переменной поперечной массы M_T . В дальнейшем планируется добавление всех имеющихся систематических погрешностей в функцию правдоподобия в виде мешающих параметров, проведение повторного фитирования и сравнение полученных результатов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ATLAS HistMaker software documentation. — URL: <https://gitlab.cern.ch/atlas-wbr-lowmu/HistMaker/-/blob/master/README.md>.
2. ATLAS WBR analysis with low mu. — URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/Vtaus13TeV>.
3. ATLAS xTauReader software documentation. — URL: https://gitlab.cern.ch/Wlep_BR/xTauReader/blob/master/doc/README.md.
4. ATLAS xTauReader software documentation. — URL: https://gitlab.cern.ch/atlas-wbr-lowmu/TRExFitter_SMWBR/-/blob/master/README.md.
5. Combined Performance (CP) Groups. — URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasPhysics>.
6. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP / S. Schael [et al.] // Phys. Rept. — 2013. — Vol. 532. — P. 119–244.
7. Event Generation with Sherpa 2.2 / E. Bothmann [et al.] // SciPost Phys. — 2019. — Vol. 7, no. 3. — P. 034.
8. Measurement of the ratio of branching fractions of W-boson decays to τ and light leptons in $t\bar{t}$ events with the ATLAS detector : тех. отч. / J. A. Mcfayden [и др.] ; CERN. — Geneva, 03.2019.
9. Monte Carlo event generators for high energy particle physics event simulation / S. Alioli [et al.]. — 2019. — Feb.
10. Multi-jet background in low-pile-up runs taken in 2017 and 2018 : тех. отч. / T. Xu [и др.] ; CERN. — Geneva, 07.2019. — ATL-COM-PHYS-2019—076.

11. Precision measurement of the mass of the τ lepton / M. Ablikim [et al.] // Phys. Rev. — 2014. — Vol. D90, no. 1. — P. 012001.
12. *Sjostrand T., Mrenna S., Skands P. Z.* A Brief Introduction to PYTHIA 8.1 // Comput. Phys. Commun. — 2008. — Vol. 178. — P. 852–867.
13. *М. Е. В.* Стандартная модель и её расширения. — НМ. : Физматлит, 2007. — 584.