

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

УДК 539.12.01

ОТЧЕТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
**АНАЛИЗ РОЖДЕНИЯ ДВУХМЮОННОГО
РЕЗОНАНСА С МАССОЙ 28 ГЭВ,
ОБНАРУЖЕННОГО НА УСТАНОВКЕ CMS, С
ИСПОЛЬЗОВАНИЕМ ДАННЫХ RUN 1 И RUN 2**

Научный руководитель

д.ф.-м.н.

Студент

_____ А. Н. Никитенко

_____ С. Д. Ильиных

Москва 2026

Содержание

Перечень сокращений и обозначений	3
Введение	4
1 Эксперимент CMS	4
1.1 Описание детектора CMS	4
1.2 Кинематика pp-столкновений и основные наблюдаемые величины	6
2 Цели, задачи и методы анализа	6
2.1 Цели и задачи	6
2.2 Используемое программное обеспечение	7
3 Реконструкция событий и критерии отбора	7
3.1 Реконструкция частиц и струй	7
3.2 В-тагирование	8
3.3 Критерии отбора событий	8
3.4 Оценка фоновых процессов	9
4 Анализ инвариантной массы и недостающего поперечного импульса	10
4.1 Метод максимального правдоподобия	10
4.2 Выбор модели для фона (F-тест)	10
4.3 Проверка стабильности фитирующей функции с использованием байес теста	12
4.4 Распределение инвариантной массы	14
4.5 Определение числа сигнальных событий с одной и двумя b-струями	15
4.6 Оптимизация отборов событий для данных Run 2 2016 года ($\sqrt{s} = 13$ ТэВ) .	16
4.7 Исследование недостающего поперечного импульса	21
4.8 Сравнение с теоретической моделью	22
5 Реализация графического интерфейса пользователя (GUI)	24
5.1 Мотивация и решенные задачи	24
5.2 Набор доступных пользователю опций	25
5.3 Техническая реализация в среде ROOT	25
Заключение	27

Перечень сокращений и обозначений

В настоящем отчете о НИР применяют следующие сокращения и обозначения:

SM Стандартная Модель физики элементарных частиц

БАК Большой адронный коллайдер

CMS Компактный мюонный соленоид (Compact Muon Solenoid)

Run 1, 2 Периоды набора данных на БАК

2HDM Двухдублетная модель Хиггса (Two-Higgs-Doublet Model)

NMSSM Минимальная суперсимметричная модель следующего поколения (Next-to-Minimal Supersymmetric Model)

p_T Поперечный импульс

η Псевдобыстрота

p_T^{miss} Недостающий поперечный импульс

OS Противоположный заряд (Opposite Sign)

Введение

Стандартная Модель (СМ) успешно описывает взаимодействия фундаментальных частиц и была подтверждена множеством экспериментов, включая открытие бозона Хиггса в 2012 году. Однако, она оставляет ряд нерешенных вопросов: природа скрытой массы, иерархия масс, барионная асимметрия, структура фермионных поколений, осцилляции нейтрино. Поэтому поиск новых легких резонансов — кандидатов на частицу скрытой массы — является актуальным направлением в современной физике частиц.

Особый интерес представляют легкие бозоны с массами 10-70 ГэВ. Их существование предсказывают расширенные хиггсовские сектора (2HDM, NMSSM модели), а также модели с новым нейтральным калибровочным Z' -бозоном, связанным с наблюдаемыми аномалиями в распадах В-мезонов. Такие частицы могут рождаться в ассоциации с b -кварками и распадаться в пару мюонов. В частности, в анализе данных CMS Run 1 и Run 2 был обнаружен избыток событий в канале распада на два мюона с массой около 28 ГэВ [1].

1 Эксперимент CMS

Эксперимент CMS (Compact Muon Solenoid — Компактный мюонный соленоид) — один из четырех основных детекторов на Большом адронном коллайдере (БАК). Детектор CMS является многоцелевым и предназначен для изучения широкого круга физических явлений, включая поиск явлений за пределами Стандартной Модели.

1.1 Описание детектора CMS

Детектор CMS [2] состоит из сверхпроводящего соленоида, создающего магнитное поле 4 Тл, трековой системы, калориметров и мюонных спектрометров. Основные компоненты детектора схематически изображены на рис. 1, 2.

Краткое описание основных систем детектора:

- **Сверхпроводящий магнит** мощностью 4 Тл. Он создает магнитное поле внутри цилиндра диаметром 6 м и длиной 12,5 м. У магнита есть возвратное ярмо, благодаря которому снаружи барреля создается сильное магнитное поле. Это позволяет детектировать отклоняющиеся мюоны. Ярмо также служит фильтром, пропуская только мюоны и слабо взаимодействующие частицы.
- **Трековая система:** Состоит из пиксельного детектора, включающего в себя три цилиндрических слоя радиусами 4, 7 и 11 см и содержащего 655 миллионов пикселей. На больших расстояниях от оси пучка (до 130 см) расположены десять слоев кремниевого стрипового детектора. Система позволяет с высокой точностью реконструировать траектории заряженных частиц и находить вершины их распадов.
- **Предливневый детектор** предназначен для регистрации высокоэнергичных фотонов. Он состоит из двух слоев свинца, между которыми расположены кремниевые сенсоры, детектирующие электрон-позитронные пары от электромагнитного ливня.

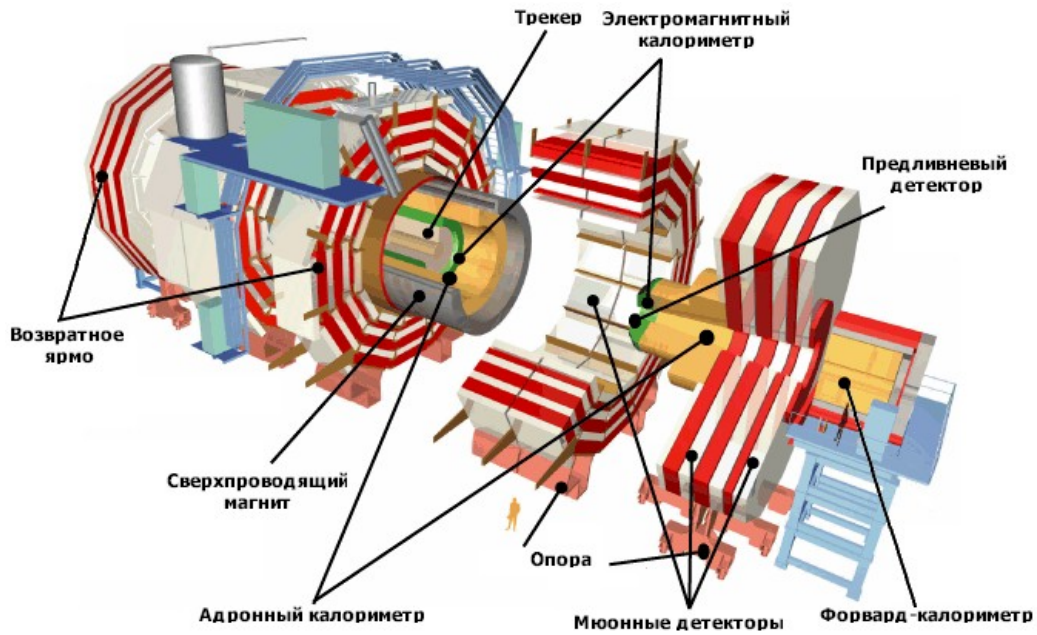


Рис. 1: Схема детектора CMS

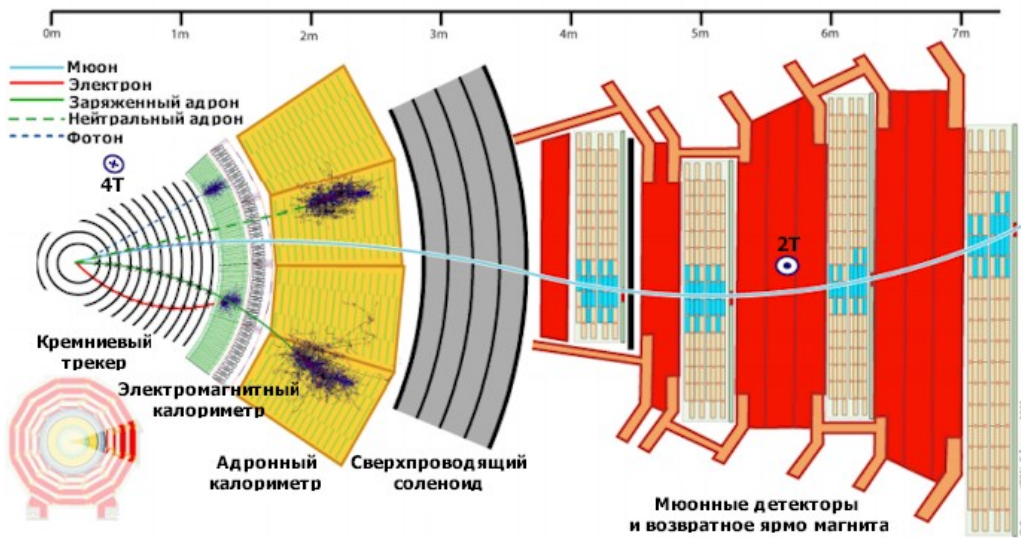


Рис. 2: Сечение детектора CMS

Предливневый детектор расположен в торцах и помогает основному электромагнитному калориметру в идентификации частиц по углу разлета фотонов.

- **Электромагнитный калориметр (ECAL):** Сделан из 76200 кристаллов вольфрамата свинца ($PbWO_4$), которые находятся в матрицах из стекловолкна. Размеры электромагнитного калориметры подобраны с учетом мольеровского радиуса вольфрамового свинца для эффективной регистрации электронов и фотонов.
- **Адронный калориметр (HCAL)** состоит из центральной, внешней, торцевой и передней секций. В центральной и торцевой секциях используется латунь. Для увеличения диапазона псевдобыстрот до 5 после мюонных детекторов установлены адронные

форвард-калориметры.

- **Мюонные спектрометры:** Расположены снаружи соленоида и используют дрейфовые трубки, катодные стриповые камеры и камеры с резистивными пластинами для идентификации и измерения импульса мюонов.

1.2 Кинематика pp-столкновений и основные наблюдаемые величины

Для описания протон-протонных столкновений в эксперименте CMS используется цилиндрическая система координат с осью z , направленной вдоль пучка. Началом отсчета считается точка столкновения двух встречных пучков. Полярный угол θ отсчитывается от положительного направления оси z , азимутальный угол ϕ определяется в плоскости Oxy . Основными наблюдаемыми величинами являются:

- **Псевдобыстрота:** $\eta = -\ln(\tan \frac{\theta}{2})$.
- **Поперечный импульс:** $p_T = p \sin \theta$.
- **Недостающий поперечный импульс p_T^{miss} :** Векторная сумма поперечных импульсов всех реконструированных частиц в событии, взятая с обратным знаком. Указывает на присутствие нейтральных, слабо взаимодействующих частиц (нейтрино, гипотетических частиц скрытой массы).
- **Инвариантная масса:** $m^2 = (\sum p_{fi})^2$, где p_{fi} — 4-импульс конечной частицы.

Подробнее о системах координат и наблюдаемых величинах в экспериментах на БАК можно прочитать в [3].

2 Цели, задачи и методы анализа

2.1 Цели и задачи

Целью данной научно-исследовательской работы является исследование событий с двухмюонным резонансом массой около 28 ГэВ в данных эксперимента CMS Run 1 и Run 2, изучение их кинематических характеристик, включая недостающий поперечный импульс, проверка устойчивости наблюдаемого избытка при различных критериях отбора и сравнение полученных распределений с феноменологической моделью рождения потенциально новой частицы.

В соответствии с целью определены следующие **задачи**:

1. Анализ критериев отбора событий для различных областей поиска (SR1, SR2 и SR2^{mod}) для данных Run 1.
2. Построение и фитирование распределения по инвариантной массе пары мюонов ($m_{\mu\mu}$) для выделения сигнальной области.

3. Проведение выбора модели фона с использованием F-теста и проверка устойчивости фитирующей функции с помощью игрушечных экспериментов.
4. Исследование распределения недостающего поперечного импульса (p_T^{miss}) в событиях с $m_{\mu\mu} \approx 28$ ГэВ.
5. Изучение вклада основных фоновых процессов (Дрелла-Яна, $t\bar{t}$, W +jets).
6. Определение числа событий с одной и двумя b-струями для области поиска $SR2^{mod}$.
7. Исследование и анализ данных Run 2 при $\sqrt{s} = 13$ ТэВ, а также оптимизация отборов по $p_T^{\mu 2}$ и $p_T^{\mu\mu}/m_{\mu\mu}$.
8. Сравнение полученных распределений p_T^{miss} с феноменологической моделью рождения потенциально новой частицы.
9. Реализация графического интерфейса пользователя для удобного задания параметров отбора и снижения риска ошибок при многократном запуске анализа.

2.2 Используемое программное обеспечение

Анализ данных проводился с использованием следующих программных инструментов:

- **ROOT:** Фреймворк для анализа данных, разработанный в ЦЕРН.
- **RooFit:** Библиотека в составе ROOT для моделирования распределений и фитирования методом максимального правдоподобия.

3 Реконструкция событий и критерии отбора

3.1 Реконструкция частиц и струй

Для реконструкции отдельных частиц в событии используется алгоритм **Particle Flow (PF)**, который комбинирует информацию со всех подсистем детектора (трекер, калориметры, мюонные спектрометры). Энергия мюонов определяется по кривизне соответствующей траектории. Энергия фотонов определяется непосредственно по измерениям из электромагнитного калориметра. Энергия электронов определяется на основе комбинации импульса электрона в вершине первичного взаимодействия, определенного трековой системой, энергии соответствующего кластера электромагнитного калориметра и суммы энергий всех фотонов тормозного излучения, пространственно совместимых с данным треком электрона. Энергия заряженных адронов определяется из комбинации их импульса, измеренного в трекере, и соответствующих энергетических потерь в калориметрах. Энергия нейтральных адронов получается из соответствующей скорректированной энергии в калориметрах. Вектор недостающего поперечного импульса в событии определяется как отрицательная векторная сумма поперечных импульсов всех кандидатов в событии:

$$p_T^{miss} = - \sum p_{Tf}$$

Для каждого события формируются адронные струи с помощью алгоритма последовательной кластеризации **Anti- k_t** [4]. Частицы внутри струй имеют небольшие различия в поперечных импульсах. Алгоритм **Anti- k_t** работает следующим образом. Определяется расстояние между двумя частицами:

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{R_{ij}^2}{R^2} \quad (1)$$

где $R = 0.5$ — параметр размера струи, а R_{ij}^2 определяется как

$$R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2. \quad (2)$$

Также определяется расстояние в импульсном пространстве между осью пучка и регистрируемой частицей:

$$d_{iB} = p_{Ti}^{-2} \quad (3)$$

Для объединения частиц в струи находят минимум набора (d_{iB}, d_{ij}) . Если d_{ij} является минимумом, частицы i, j объединяются в одну и удаляются из списка частиц. Если d_{iB} является минимумом, частица i помечается как финальная струя и удаляется из списка. Итерации повторяются, пока все частицы не будут отнесены к одной из реконструированных струй.

3.2 В-тагирование

Струи, происходящие от b -кварков, идентифицируются (b -тагируются) с помощью алгоритмов многомерного анализа (MVA). Адроны, содержащие b -кварк, имеют время жизни порядка пикосекунды (характеристическая длина распада $c\tau \approx 450 - 500$ мкм). Поэтому с помощью трековой системы можно обнаружить вторичную вершину, в которой адроны распадаются на заряженные частицы. Таким образом, струя, содержащая реконструированную вторичную вершину, скорее всего ассоциирована с b -кварком. В используемом алгоритме эффективность маркировки струй с b -кварком составляет примерно 42% при 0,05% частоте ошибочной идентификации струй с легкими кварками или глюоном [5].

3.3 Критерии отбора событий

Для подавления фоновых процессов (*Дрелла-Ян*, $t\bar{t}$, $W + \text{jets}$) и выделения сигнальных событий с резонансным рождением пары мюонов применяются строгие критерии отбора. Основные требования для отборов Run 1 приведены в таблице 1.

Критерий	SR1	SR2	SR2 ^{mod}
Мюоны		OS, $p_T > 25$ ГэВ, $ \eta < 2.1$	OS, $p_T > 31, 26$ ГэВ, $ \eta < 2.1$
$m_{\mu\mu}$		> 12 ГэВ	
Струи		$p_T > 30$ ГэВ	
b-струя		$ \eta \leq 2.4$	
Доп. струи $ \eta \leq 2.4$	только 1 центральная струя (b-струя)	ровно 2 центральные струи (хотя бы одна b-струя)	ровно 2 центральные струи (хотя бы одна b-струя)
Доп. струи $2.4 < \eta < 4.7$	≥ 1 струя	0 струй	0 струй
p_T^{miss}	–	< 40 ГэВ	–
$\Delta\phi(\mu\mu, jj)$	–	> 2.5 рад.	–
$\frac{p_T}{m_{\mu\mu}}$	–	–	> 2.0

Таблица 1: Критерии отбора событий в трех областях поиска Run 1.

3.4 Оценка фоновых процессов

Для проверки состоятельности модели фона и понимания структуры событий коллаборацией CMS построены распределения инвариантной массы пары мюонов для основных фоновых процессов: Дрелла-Яна (DY), $t\bar{t}$ и образование пары бозонов (VV). На рис. 3 показаны вклады этих процессов в области поиска SR1 и SR2 из статьи [1].

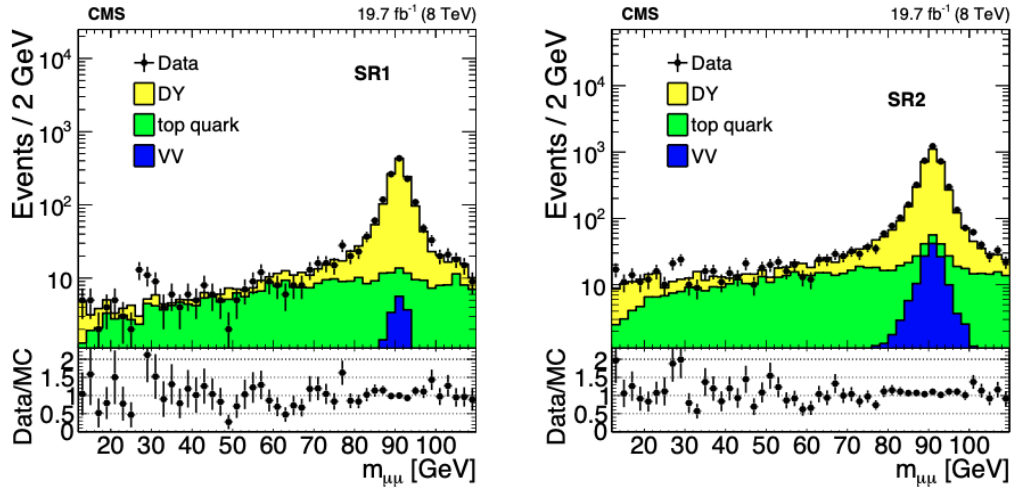


Рис. 3: Распределение инвариантной массы пары мюонов $m_{\mu\mu}$ в SR1 (слева) и SR2 (справа) при энергии 8 ТэВ, с ожидаемыми фоновыми процессами, основанными на моделировании.

Как видно из графиков, доминирующим фоном в исследуемой области масс является процесс образования пары $t\bar{t}$ -кварков. Процессы Дрелла-Яна и образование пары бозонов, хотя и подавлены критериями отбора (требование двух струй, одна из которых b-тагирована, и отсутствие дополнительных струй), вносят заметный вклад в область низких масс (< 20 ГэВ) и могут влиять на оценку формы фона под сигналом.

4 Анализ инвариантной массы и недостающего поперечного импульса

4.1 Метод максимального правдоподобия

Спектр инвариантной массы пары мюонов фитируется с помощью свертки функций Брейта-Вигнера и Гаусса для моделирования возможного сигнала, где наблюдается избыток событий. Функция Брейта-Вигнера описывает форму резонанса, а распределение Гаусса имитирует экспериментальное массовое разрешение 0,45 ГэВ для системы пары мюонов с массой 28 ГэВ. Для фитирования фона используется полином второй степени, параметры которого могут свободно варьироваться при фитировании. Чтобы количественно охарактеризовать любой потенциальный избыток событий, выполняется фитирование распределения массы димюонов в диапазоне 12–70 ГэВ методом максимального правдоподобия, используя следующее выражение для правдоподобия:

$$\mathcal{L}(m_X, \Gamma_{\mu\mu}, a_1, a_2) = \frac{(N_S + N_B)^N}{N!} e^{-(N_S + N_B)} \times \prod_{i=1}^N \left[\frac{N_S}{N_S + N_B} p_i^S(m_X, \Gamma_{\mu\mu}) + \frac{N_B}{N_S + N_B} p_i^B(a_1, a_2) \right], \quad (4)$$

где N — число наблюдаемых событий, N_S — число сигнальных событий, N_B — число фоновых событий, а p_i^S и p_i^B — функции плотности вероятности для сигнала и фона соответственно.

Сигнальная компонента описывается сверткой Брейта-Вигнера с Гауссом:

$$p^S(m_X, \Gamma_{\mu\mu}) = \int \frac{\Gamma_{\mu\mu}/2\pi}{(m' - m_X)^2 + (\Gamma_{\mu\mu}/2)^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(m_X - m')^2}{2\sigma^2}} dm' \quad (5)$$

где m_X и $\Gamma_{\mu\mu}$ — параметры резонанса.

Фон моделируется полиномом второй степени:

$$p^B(\vec{a}) = 1 + \sum_{k=1}^2 a_k \left(\frac{m - m_{\min}}{m_{\max} - m_{\min}} \right)^k \quad (6)$$

где $m_{\min} = 12$ ГэВ и $m_{\max} = 70$ ГэВ — границы исследуемого диапазона масс, а $\vec{a} = (a_1, a_2)$ — коэффициенты полинома.

4.2 Выбор модели для фона (F-тест)

Для обоснованного выбора степени полинома, описывающего фон, был применен F-тест, сравнивающий конкурирующие модели.

- H_0 (фон): $P_B(m) = \mathcal{P}_n(m; \vec{a})$
- H_1 (фон): $P_B(m) = \mathcal{P}_{n+1}(m; \vec{a})$

Результаты теста (на основе минимума логарифма правдоподобия $-\ln \mathcal{L}$ только от фона) представлены в таблице 2.

Степень полинома	SR1		SR2	
	$-\ln \mathcal{L}$	$\Delta(-\ln \mathcal{L})$	$-\ln \mathcal{L}$	$\Delta(-\ln \mathcal{L})$
1	23,92	–	480,96	–
2	25,29	1,37	484,22	3,26
3	25,43	0,14	484,22	0,00
4	25,62	0,19	485,23	1,01

Таблица 2: Результаты F-теста для выбора порядка полиномиальной модели фона в двух областях поиска. Критерий значимости: $\Delta(\ln \mathcal{L}) > 0,5$, что соответствует одному стандартному отклонению.

На рис. 4, 5 представлены результаты фитирования распределения инвариантной массы димюонов полиномами 1-й, 2-й, 3-й и 4-й степени.

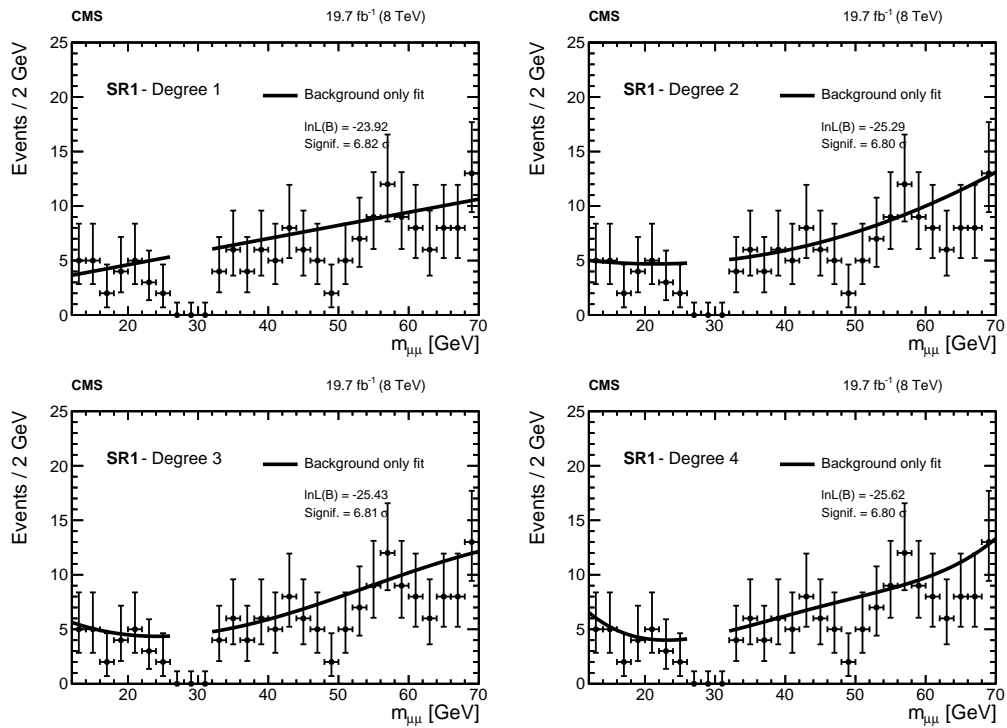


Рис. 4: Фитирование распределения инвариантной массы димюонов полиномами различной степени для SR1. Точки — данные, линии — результаты фитирования.

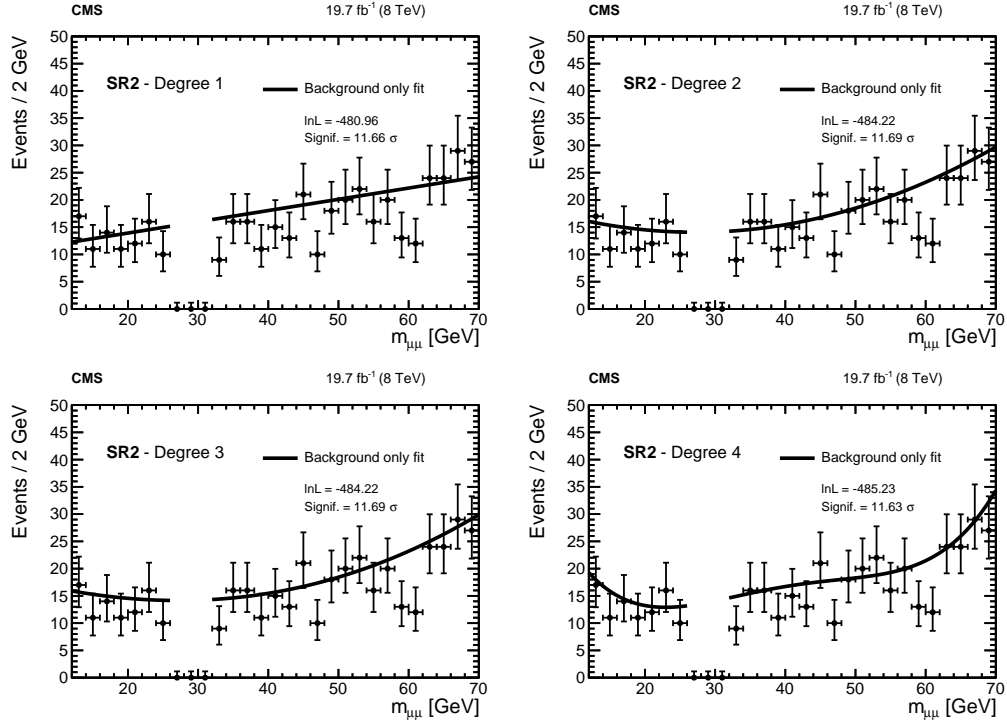


Рис. 5: Фитирование распределения инвариантной массы димюонов полиномами различной степени для SR2. Точки — данные, линии — результаты фитирования.

Основываясь на значительном улучшении при переходе от полинома 1-й степени к полиному 2-й степени и незначительных улучшениях для высших степеней (для обеих областей поиска), для фитирования фона была выбрана модель полинома 2-й степени.

4.3 Проверка стабильности фитирующей функции с использованием байес теста

Для оценки систематической неопределенности, связанной с выбором параметров и формы распределения фона, был проведен байес тест с использованием игрушечных экспериментов. В рамках этого подхода рассматривались четыре различные комбинации моделей фона: полином Чебышева второй степени (*Chebyshev(2)*) и экспоненциальная функция от полинома второй степени ($\exp(a_0 + a_1 \cdot x + a_2 \cdot x^2)$) как для сигнальной области, так и для сайдбендов. Анализ выполнялся независимо для двух областей отбора: SR1 и SR2.

Для каждой комбинации моделей фона было смоделировано 5000 событий. В каждом эксперименте генерировались события с известными параметрами сигнала (истинное число событий $N_{sig} = 23$, истинная масса $m = 28.0$ ГэВ, ширина $width = 1.4$ ГэВ) и фона, после чего выполнялся фит для восстановления этих параметров. Для оценки качества восстановления и наличия систематического смещения анализировались распределения pull-величин:

$$pull(N_{sig}) = \frac{N_{sig}^{fit} - N_{sig}^{true}}{\sigma_{N_{sig}}}, \quad pull(m) = \frac{m^{fit} - m^{true}}{\sigma_m} \quad (7)$$

В идеальном случае эти распределения должны иметь форму стандартного нормального распределения, то есть $\mu = 0$ и $\sigma = 1$. Пример распределения pull-величин для области поиска SR1 и конфигурации "моделирование распределением полином Чебышева II степени - фитирование распределением экспоненциальная функция от полинома II степени" приведен на рис. 6. Критерием приемлемости служит отклонение среднего значения pull-распределения от 0 не более чем на 20% от ожидаемого среднего, т.е. $|\mu_{pull}| < 0.2$. В таблицах 3 и 4 приведен результат байес теста для SR1 и SR2 соответственно.

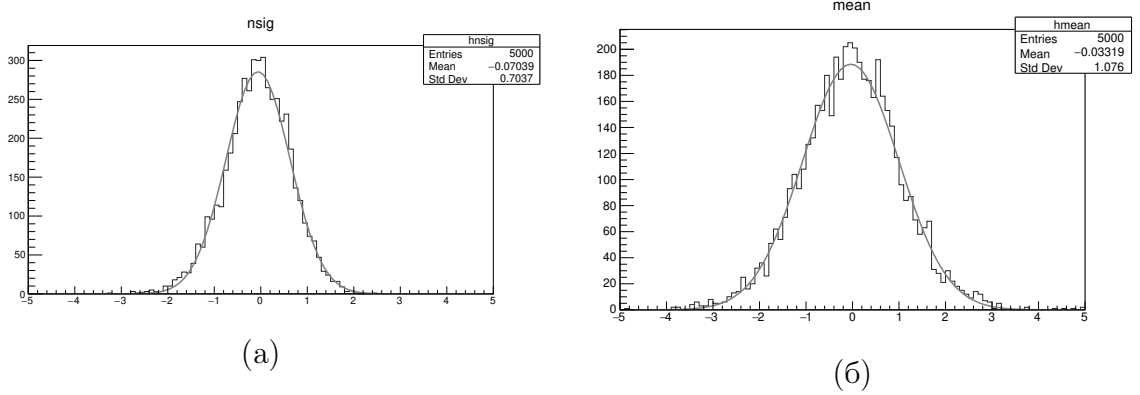


Рис. 6: Распределение pull-величин (количество сигнальных событий N_{sig} и масса m).

Модель фона (SR / CR)	N_{sig}	m , ГэВ	$\mu_{pull}(N_{sig})$	$\mu_{pull}(m)$
<i>Chebyshev(2) / Chebyshev(2)</i>	24.23 ± 7.04	28.84 ± 0.41	0.0217	-0.0269
<i>Chebyshev(2) / exp(P_2)</i>	25.35 ± 6.63	28.84 ± 0.40	-0.0704	-0.0332
<i>exp(P_2) / Chebyshev(2)</i>	20.75 ± 6.50	28.62 ± 0.38	0.0861	-0.0244
<i>exp(P_2) / exp(P_2)</i>	21.10 ± 6.49	28.62 ± 0.38	0.0194	-0.0288

Таблица 3: Результаты фитирования и средние значения распределений pull-величин для сигнальной области SR1 (5000 событий, $\sqrt{S} = 8$ ТэВ).

Модель фона (SR / CR)	N_{sig}	m , ГэВ	$\mu_{pull}(N_{sig})$	$\mu_{pull}(m)$
<i>Chebyshev(2) / Chebyshev(2)</i>	24.33 ± 8.34	27.96 ± 0.33	0.0771	-0.0086
<i>Chebyshev(2) / exp(P_2)</i>	23.80 ± 7.82	27.95 ± 0.33	-0.0508	-0.0168
<i>exp(P_2) / Chebyshev(2)</i>	19.85 ± 8.25	28.06 ± 0.50	0.1932	0.0626
<i>exp(P_2) / exp(P_2)</i>	19.77 ± 7.68	28.05 ± 0.50	0.0752	0.0442

Таблица 4: Результаты фитирования и средние значения распределений pull-величин для сигнальной области SR2 (5000 событий, $\sqrt{S} = 8$ ТэВ).

Как видно из таблиц 3 и 4, для всех рассмотренных комбинаций моделей фона средние значения pull-распределений находятся в пределах $|\mu_{pull}| < 0.2$ для обоих параметров (инвариантной массы двух мюонов и числа сигнальных событий), что свидетельствует об отсутствии систематического смещения при определении положения сигнального пика.

4.4 Распределение инвариантной массы

На рис. 7 представлено распределение инвариантной массы пары мюонов, полученное для данных Run 1 в двух областях поиска SR1 и SR2. Фитирование выполнено в диапазоне 12–70 ГэВ. В области 25–30 ГэВ наблюдается избыток событий над фоном.

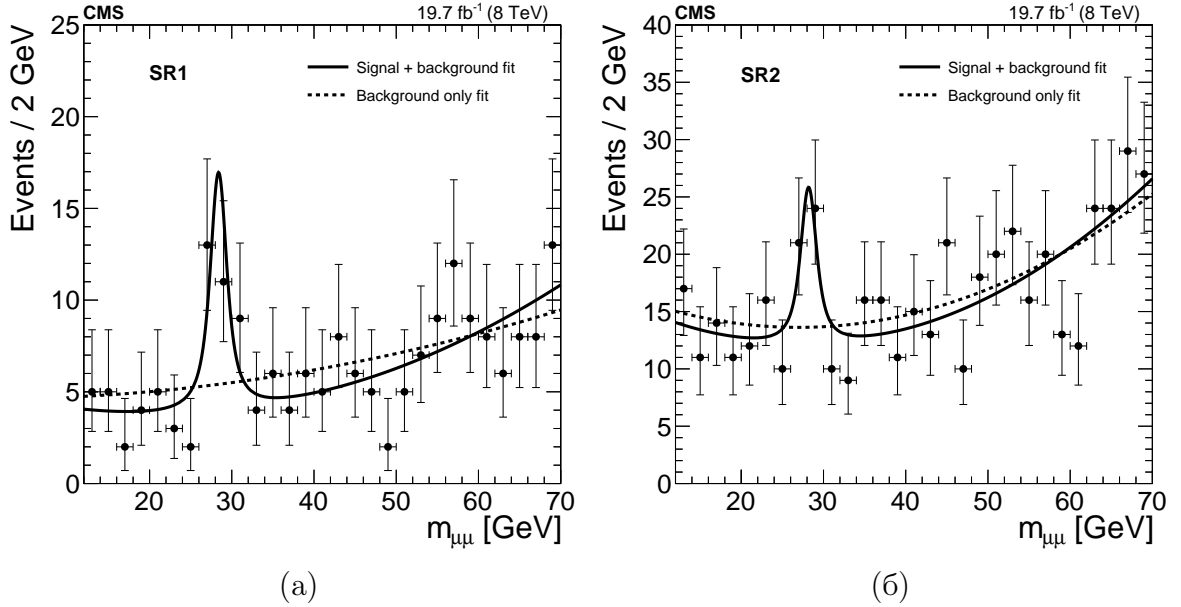


Рис. 7: Распределение инвариантной массы пары мюонов для SR1 (а) и SR2 (б).

Для модифицированной области поиска $SR2^{mod}$ также построено распределение по инвариантной массе $m_{\mu\mu}$ — см. рис. 8.

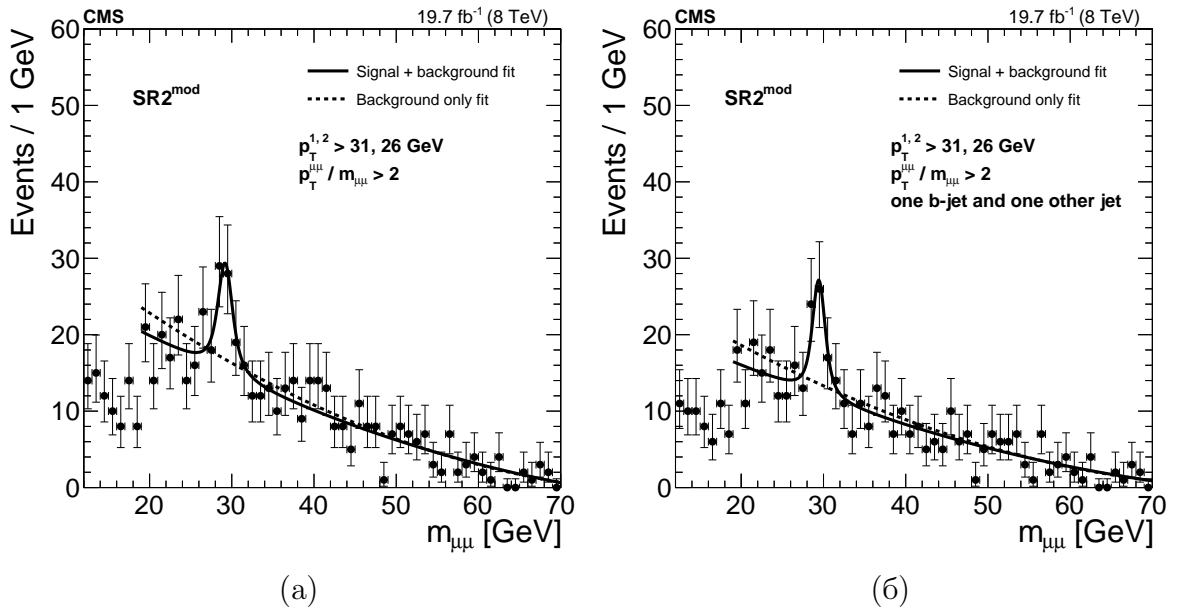


Рис. 8: Распределение инвариантной массы пары мюонов для $SR2^{mod}$ в случае отборов согласно таблице 1 (а) и в случае с одной b-струей (б).

4.5 Определение числа сигнальных событий с одной и двумя b -струями

Для построения теоретической модели необходимо определить истинное число событий с дополнительной струей, ассоциированной с b -кварком, и с дополнительной "легкой" струей (содержащей глюон или легкий кварк). Для распределения на рис. 8 (а) получено 44 сигнальных события при условии, что струи две и одна из них точно содержит b -кварк. Для распределения на рис. 8 (б) получено 39 сигнальных событий при условии, что одна струя b -тагирована, а вторая - любая другая (содержит глюон или легкий кварк). Эффективность b -тагирования $\epsilon_b = 0,42$, вероятность ложного тагирования струи как содержащей b -кварк $f_{mistag} = 0,0005$ (см. [1]).

Тогда для истинной bl -пары:

- Вероятность того, что для двух струй первая правильно помечается как b -струя, а вторая ложно помечается как b -струя: $P(n = 2) = \epsilon_b \cdot f_{mistag}$.
- Вероятность того, что b -струя помечается, а вторая не помечается: $P = \epsilon_b \cdot (1 - f_{mistag})$. Вероятность того, что b -струя не помечается, а вторая ложно помечается как b -струя: $P = (1 - \epsilon_b) \cdot f_{mistag}$. Тогда $P(n = 1) = \epsilon_b \cdot (1 - f_{mistag}) + (1 - \epsilon_b) \cdot f_{mistag}$
- Вероятность того, что обе струи не помечаются как b -струи: $P(n = 0) = (1 - \epsilon_b) \cdot (1 - f_{mistag})$

Для нашего отбора (две струи, хотя бы одна из них b -струя) вероятность получается равной:

$$P_1(bl) = 1 - P(n = 0) = 1 - (1 - \epsilon_b) \cdot (1 - f_{mistag}) \quad (8)$$

Аналогично для истинной bb -пары:

- $P(n = 2) = \epsilon_b^2$.
- $P(n = 1) = 2 \cdot \epsilon_b \cdot (1 - \epsilon_b)$.
- $P(n = 0) = (1 - \epsilon_b)^2$

Для нашего отбора (две струи, хотя бы одна из них b -струя) вероятность получается равной:

$$P_1(bb) = 1 - P(n = 0) = 1 - (1 - \epsilon_b)^2 = 2\epsilon_b - \epsilon_b^2 \quad (9)$$

Аналогично для нового отбора (одна b -струя и одна другая струя):

$$P_2(bl) = \epsilon_b(1 - f_{mistag}) + (1 - \epsilon_b)f_{mistag} \quad (10)$$

$$P_2(bb) = 0 \quad (11)$$

Таким образом, имеем систему уравнений (для фиксированной массы и ширины распределения):

$$\begin{cases} N_{bl}P_1(bl) + N_{bb}P_1(bb) = 44 \pm 13, \\ N_{bl}P_2(bl) + N_{bb}P_2(bb) = 39 \pm 11, \end{cases} \quad (12)$$

где N_{bl} и N_{bb} - истинное число событий с одним и двумя b-кварками соответственно. Решая эту систему уравнений, получаем:

$$\begin{cases} N_{bl} \approx 93 \pm 26, \\ N_{bb} \approx 8 \pm 28. \end{cases} \quad (13)$$

Полученные значения N_{bl} и N_{bb} представляют собой истинное количество событий в анализируемой выборке. Однако для построения теоретической модели необходимо знать, сколько событий каждого типа реально наблюдается детектором после применения алгоритма b-тагирования, то есть какой вклад вносят события bl и bb в отобранные 44 события.

Вклад событий с одним b-кварком в первую выборку составляет:

$$N_{bl}^{obs} = N_{bl} \cdot P_1(bl) = (93 \pm 26) \cdot 0,420 \approx 39 \pm 11. \quad (14)$$

Вклад событий с двумя b-кварками в первую выборку составляет:

$$N_{bb}^{obs} = N_{bb} \cdot P_1(bb) = (7,5 \pm 28) \cdot 0,664 \approx 5 \pm 19. \quad (15)$$

Сумма вкладов $39 + 5 = 44$ соответствует полному числу отобранных событий в сигнальной области.

4.6 Оптимизация отборов событий для данных Run 2 2016 года ($\sqrt{s} = 13$ ТэВ)

Для проверки воспроизводимости наблюдаемого избытка событий в области $m_{\mu\mu} \approx 28$ ГэВ был выполнен дополнительный анализ данных Run 2 при энергии столкновений $\sqrt{s} = 13$ ТэВ и интегральной светимости $35,9 \text{ фб}^{-1}$. Анализ проводился для области поиска SR1. Особое внимание уделялось зависимости величины сигнала от кинематических требований на второй мюон и на буст димюонной системы, определяемый как

$$boost = \frac{p_T^{\mu\mu}}{m_{\mu\mu}}. \quad (16)$$

На рис. 9 представлено распределение по величине $p_T^{\mu\mu}/m_{\mu\mu}$ при отборе $p_T^{\mu 1,2} > 25$ ГэВ. Видно, что для событий сигнальной области значения $p_T^{\mu\mu}/m_{\mu\mu}$ лежат преимущественно в диапазоне от 1,5 до 4,0. Поэтому далее рассмотрен отбор по бусту димюонной системы как дополнительная кинематическая переменная для повышения статистической значимости сигнала.

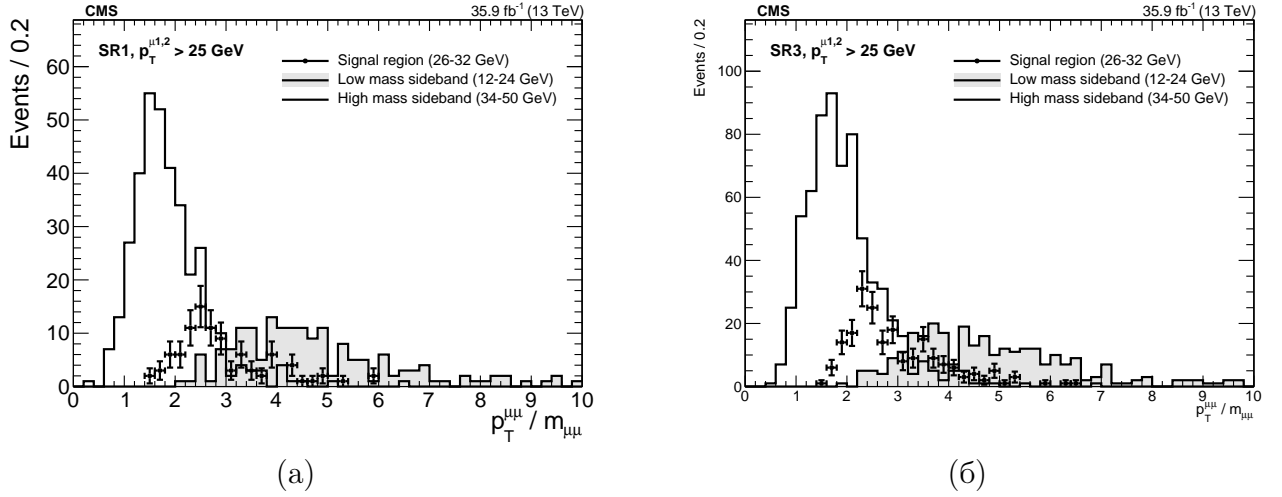


Рис. 9: Распределение по величине $p_T^{\mu\mu} / m_{\mu\mu}$ для данных Run 2 при $\sqrt{s} = 13$ ТэВ в областях поиска SR1 (а) и SR3 (б) при требовании $p_T^{\mu_{1,2}} > 25$ ГэВ.

На рис. 10 приведено распределение поперечного импульса второго мюона $p_T^{\mu_2}$ при отсутствии дополнительного ограничения на буст ($p_T^{\mu\mu} / m_{\mu\mu} > 0$). Наблюдается пик событий в сигнальной области (26-32 ГэВ) при $p_T^{\mu_2} \approx 17 - 18$ ГэВ.

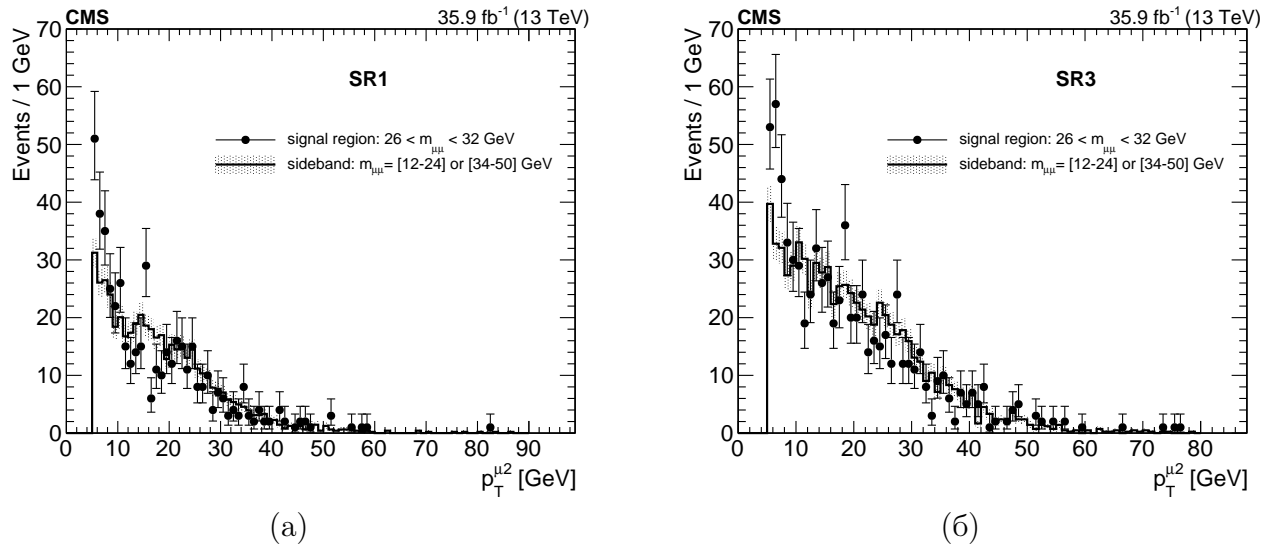
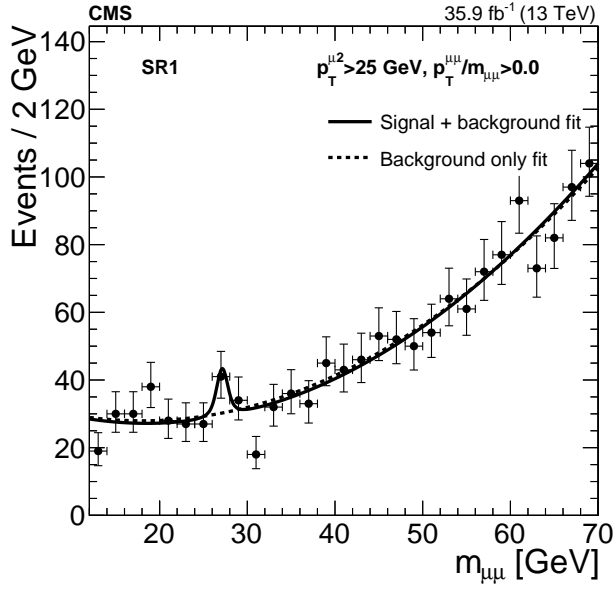
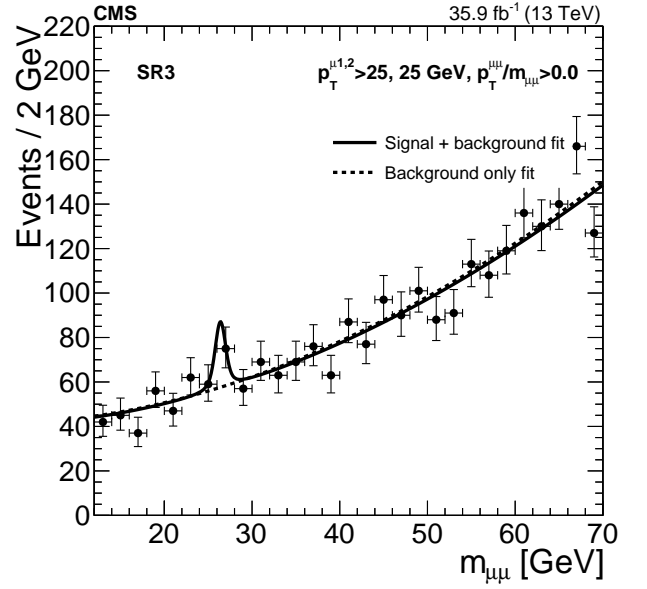


Рис. 10: Распределение поперечного импульса второго мюона $p_T^{\mu_2}$ для данных Run 2 при $\sqrt{s} = 13$ ТэВ в области поиска SR1 (а) и SR3 (б).

До оптимизации было построено распределение инвариантной массы пары мюонов при базовом отборе $p_T^{\mu_2} > 25$ ГэВ и $p_T^{\mu\mu} / m_{\mu\mu} > 0$ (рис. 11), соответствующем отбору SR1 Run 1. Статистическая значимость избытка событий в области массы $m_{\mu\mu} \approx 28$ ГэВ при таком отборе составляет $Z = 2.0$ для SR1 и $Z = 2.3$ для SR3. Это связано с возросшим фоном от $t\bar{t}$ при увеличении энергии столкновения протонов, поэтому было принято решение оптимизировать отборы.



(a)



(б)

Рис. 11: Распределение инвариантной массы пары мюонов до оптимизации для данных Run 2 при $\sqrt{s} = 13$ ТэВ в области поиска SR1 (а) и SR3 (б).

Для выбора оптимальных значений порогов выполнен двумерный скан по параметрам $p_T^{\mu 2}$ и $p_T^{\mu\mu} / m_{\mu\mu}$. Для каждой пары значений проводилось фитирование распределения инвариантной массы пары мюонов для фоновой (полином II степени для SR1 и полином III степени для SR3) и сигнальной (свертка Брейта-Вигнера и Гаусса) компоненты. Статистическая значимость оценивалась через изменение логарифма правдоподобия. Результаты представлены в виде тепловой карты на рис. 12 и 13.

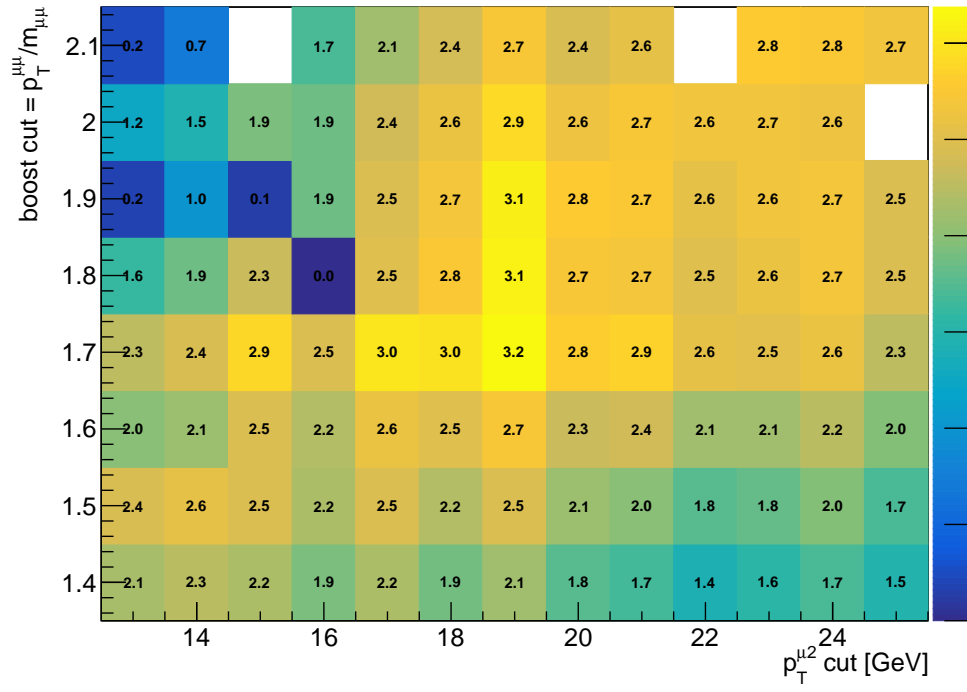


Рис. 12: Тепловая карта статистической значимости для данных Run 2 при $\sqrt{s} = 13$ ТэВ в зависимости от порогов на $p_T^{\mu 2}$ и $p_T^{\mu\mu}/m_{\mu\mu}$. Область поиска SR1.

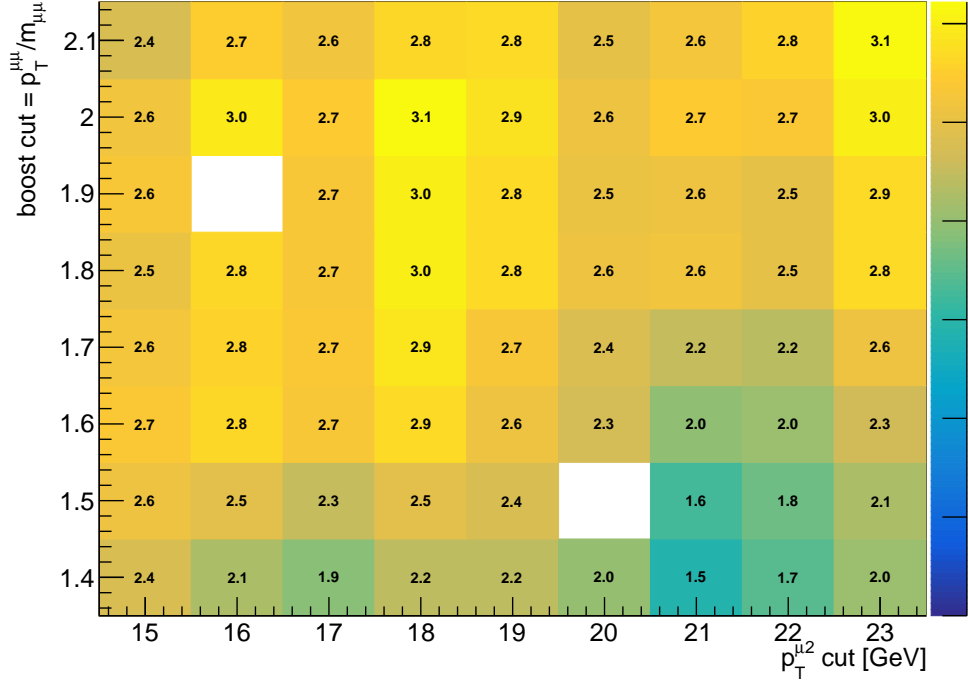


Рис. 13: Тепловая карта статистической значимости для данных Run 2 при $\sqrt{s} = 13$ ТэВ в зависимости от порогов на $p_T^{\mu 2}$ и $p_T^{\mu\mu}/m_{\mu\mu}$. Область поиска SR3.

Для иллюстрации влияния порога на буст на статистическую значимость рассмотрены три значения порога на буст: $p_T^{\mu\mu}/m_{\mu\mu} > 1.4$ — нижний порог отбора по бусту, 1.7 — оптимальное значение и 2.1 — верхний порог — при условии $p_T^{\mu 1,2} > 25, 19$ ГэВ. Соответствующие распределения инвариантной массы приведены на рис. 14.

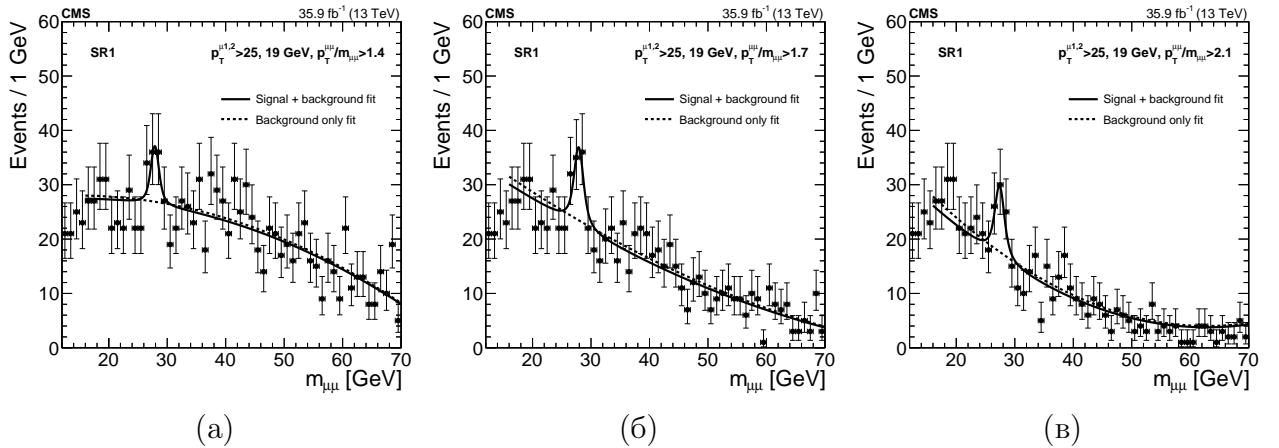


Рис. 14: Распределения инвариантной массы пары мюонов для различных порогов на буст. Область поиска SR1.

Аналогичные распределения построены и для области поиска SR3: $p_T^{\mu\mu}/m_{\mu\mu} > 1.4$ — нижний порог отбора по бусту, 1.7 — оптимальное значение и 2.1 — верхний порог — при

условии $p_T^{\mu_{1,2}} > 25, 18$ ГэВ. Соответствующие распределения инвариантной массы приведены на рис. 15.

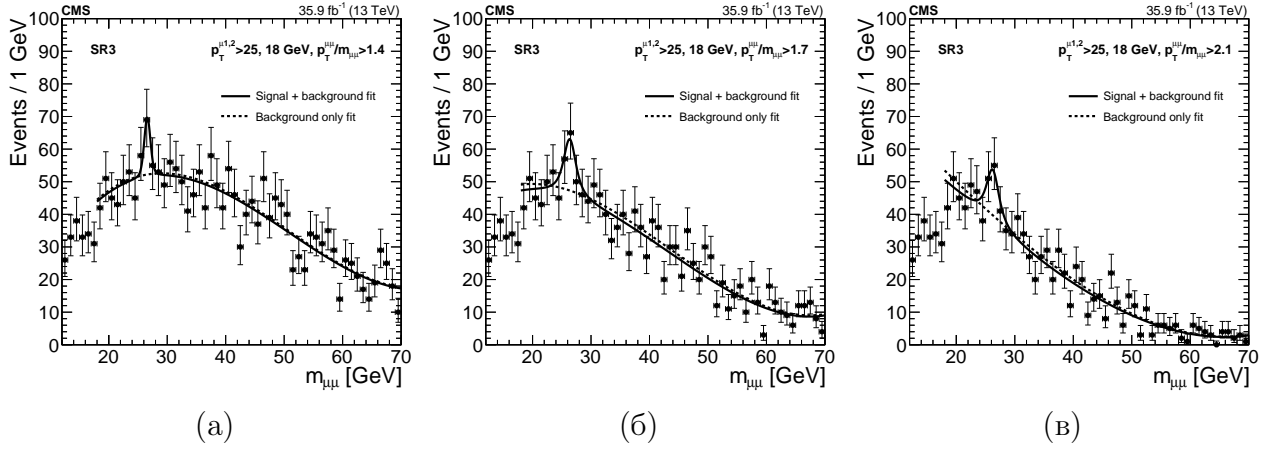


Рис. 15: Распределения инвариантной массы пары мюонов для различных порогов на буст. Область поиска SR3.

Количественные результаты фитирования приведены в таблице 5.

Область поиска	Порог на $p_T^{\mu_2}$	Порог на буст $p_T^{\mu\mu}/m_{\mu\mu}$	N_{sig}	Z
SR1	$p_T^{\mu_2} > 19$	$p_T^{\mu\mu}/m_{\mu\mu} > 1.4$	17.9	2.1
SR1	$p_T^{\mu_2} > 19$	$p_T^{\mu\mu}/m_{\mu\mu} > 1.7$	29.3	3.2
SR1	$p_T^{\mu_2} > 19$	$p_T^{\mu\mu}/m_{\mu\mu} > 2.1$	24.4	2.7
SR3	$p_T^{\mu_2} > 18$	$p_T^{\mu\mu}/m_{\mu\mu} > 1.4$	26.1	2.2
SR3	$p_T^{\mu_2} > 18$	$p_T^{\mu\mu}/m_{\mu\mu} > 1.7$	41.2	2.9
SR3	$p_T^{\mu_2} > 18$	$p_T^{\mu\mu}/m_{\mu\mu} > 2.1$	40.2	2.8

Таблица 5: Число сигнальных событий и статистическая значимость для различных порогов на буст.

Из таблицы 5 следует, что максимальная статистическая значимость достигается при значении $p_T^{\mu\mu}/m_{\mu\mu} > 1.7$. При дальнейшем увеличении порога происходит уменьшение статистики, что приводит к снижению чувствительности анализа. Таким образом, в качестве оптимального отбора для данных Run 2 в области поиска SR1 выбраны отборы $p_T^{\mu_1} > 25$ ГэВ, $p_T^{\mu_2} > 19$ ГэВ и $p_T^{\mu\mu}/m_{\mu\mu} > 1.7$. Для области поиска SR3 выбраны отборы $p_T^{\mu_1} > 25$ ГэВ, $p_T^{\mu_2} > 18$ ГэВ и $p_T^{\mu\mu}/m_{\mu\mu} > 1.7$.

4.7 Исследование недостающего поперечного импульса

Одним из ключевых аспектов данной работы является изучение распределения недостающего поперечного импульса (p_T^{miss}) в событиях, попавших в сигнальную область масс ($26 < m_{\mu\mu} < 32$ ГэВ). Наличие ненулевого p_T^{miss} может указывать на рождение в событии частиц, не оставляющих сигнала в детекторе.

На рис. 16 представлены распределения p_T^{miss} для событий сигнальной области, полученные для данных Run 1 и Run 2 (энергия 13 ТэВ в системе центра масс, интегральная светимость 35.9 fb^{-1}).

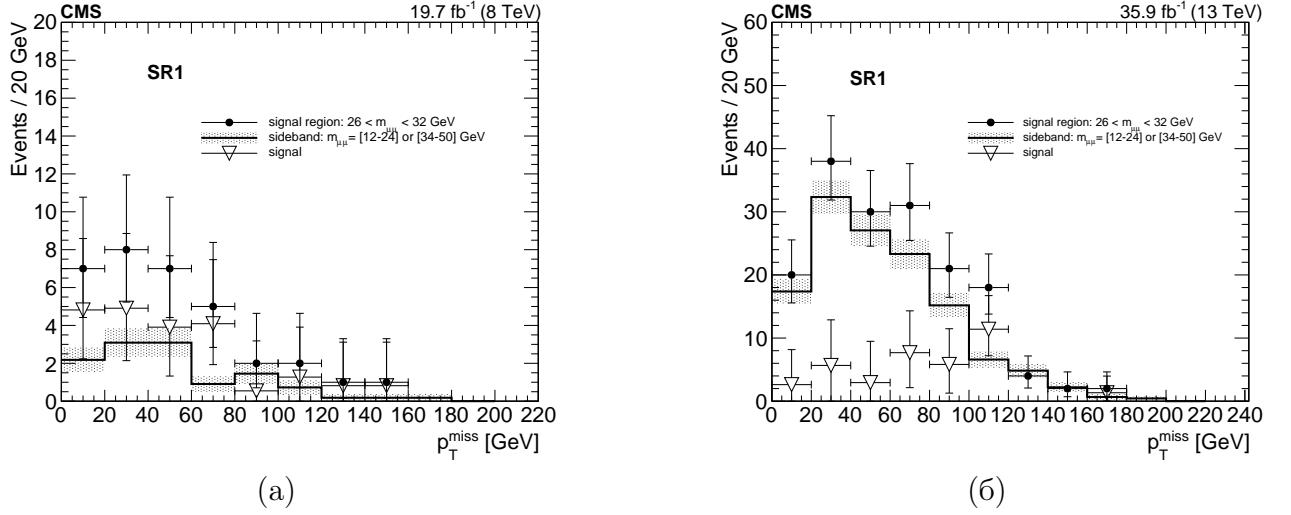


Рис. 16: Распределение недостающего поперечного импульса (p_T^{miss}) в событиях с $m_{\mu\mu} \approx 28$ ГэВ для данных (а) Run 1 (8 ТэВ) и (б) Run 2 (13 ТэВ) в области поиска SR1.

Анализ распределений p_T^{miss} показывает, что в событиях с инвариантной массой $m_{\mu\mu}$ около 28 ГэВ наблюдается ненулевая недостающая энергия. Это отличает данные события от чистого фона Дрелла-Яна, для которого p_T^{miss} в основном сосредоточена вблизи нуля. Наличие p_T^{miss} , а также требования на ассоциированные b -струи, согласуется с гипотезой о рождении резонанса в ассоциации с b -кварками и последующим распадом с участием нейтральных, слабо взаимодействующих частиц.

4.8 Сравнение с теоретической моделью

В качестве рабочей теоретической модели рассматривается резонансная модель, в которой в протон–протонных столкновениях рождается тяжелое состояние Y_S (скаляр) с массой порядка 250 ГэВ в ассоциации с b -кварками и струей, после чего оно распадается на два состояния с массой около 28 ГэВ. Одно из них, $X_V(28)$ - виртуальное состояние (продольно поляризованный вектор), распадается в пару мюонов, а второе — в фермионы тёмного сектора dm_f с массой 70 ГэВ каждый, что приводит к конечному состоянию с $\mu^+\mu^-$, b -струями и недостающим поперечным импульсом. Диаграммы Фейнмана процесса рождения пары мюонов показаны на рис. 17.

$$pp \rightarrow Y_S(250) b\bar{b}j, \quad Y_S(250) \rightarrow X_V(28) X_V^*(28), \quad X_V(28) \rightarrow \mu^+\mu^-, \quad X_V^*(28) \rightarrow dm_f(70) dm_f(70).$$

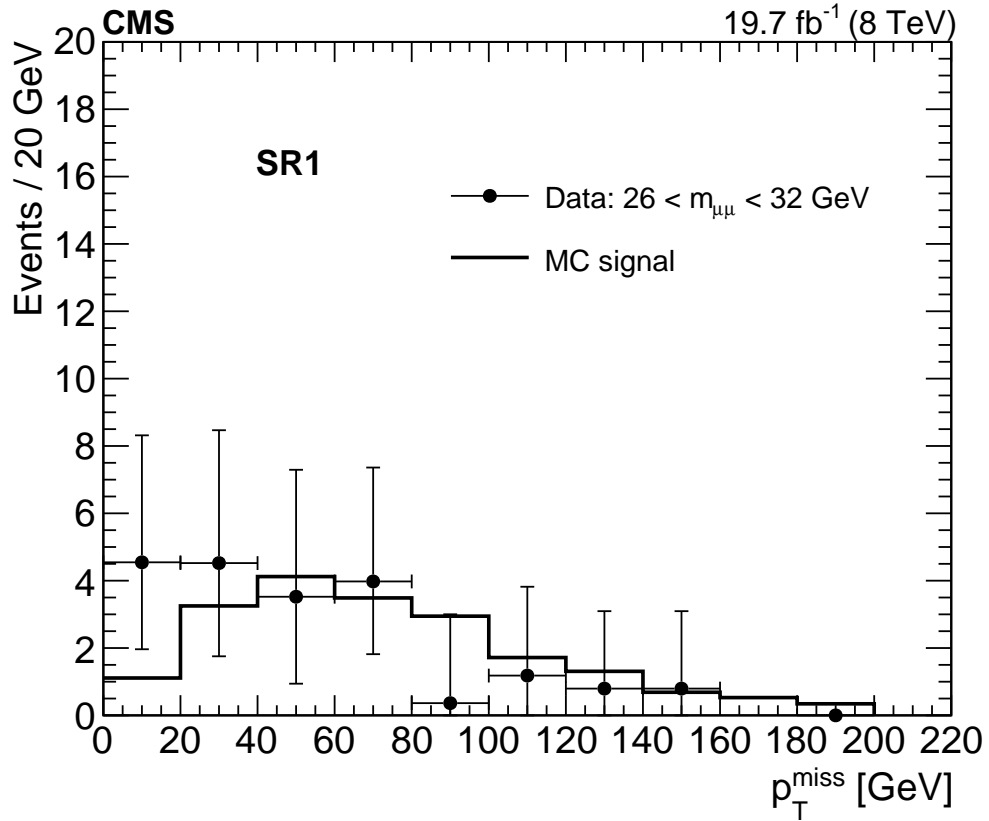


Рис. 18: Распределение недостающего поперечного импульса для Run 1 (8 ТэВ) и области поиска SR1. Точки — события в сигнальной области, сплошная линия — смоделированные данные.

5 Реализация графического интерфейса пользователя (GUI)

5.1 Мотивация и решенные задачи

В ходе анализа данных Run 1 и Run 2 с различными критериями отбора (SR1, SR2, SR2^{mod}) существенная часть времени тратится на ручное изменение параметров отбора, входных и выходных файлов. При таком подходе повышается риск ошибок пользователя. Для решения данной проблемы реализован графический интерфейс пользователя (Graphical User Interface - GUI), интегрированный в текущий скрипт для отбора нужных событий.

Интерфейс решает следующие задачи:

1. Стандартизация ввода параметров. Пользователь вводит параметры в структурированном виде, где каждому параметру соответствует отдельное поле. Это снижает вероятность неоднозначной интерпретации значений и облегчает проверку параметров перед запуском.
2. Повышение удобства и читаемости конфигурации. Группировка по смысловым блокам (например, параметры отбора мюонов, пары мюонов, струй, b -тагирование, фоновые

настройки и т.д.) обеспечивает быстрое переключение между секциями и снижает риск человеческого фактора при работе с большим числом параметров.

5.2 Набор доступных пользователю опций

На рис. 19 представлено окно графического интерфейса. Реализована возможность ручного ввода значений или управление ими с помощью спинеров.

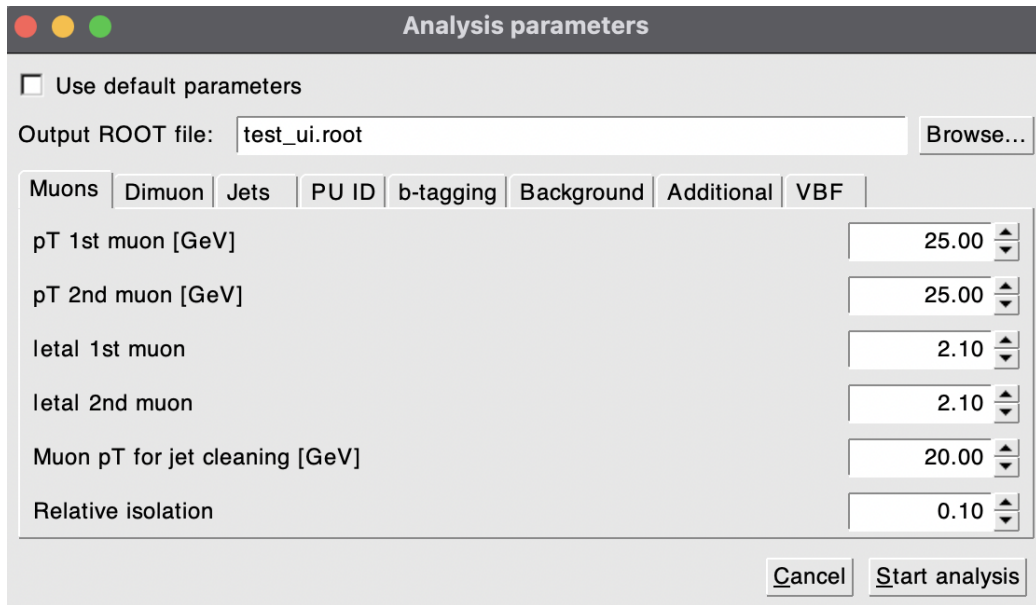


Рис. 19: GUI. Первая вкладка с настройкой параметров.

В рамках GUI предоставляются следующие возможности конфигурирования:

- Выбор режима параметров: использование набора параметров по умолчанию или ручной ввод значений. При выборе первого варианта, изменение любых параметров становится невозможным.
- Ввод численных параметров через текстовые поля ввода с фиксированным форматом.
- Логическая навигация по разделам: список параметров разбит на логические вкладки. Это исключает необходимость прокручивать длинный “плоский” список параметров и обеспечивает устойчивую структуру интерфейса при расширении набора переменных.

5.3 Техническая реализация в среде ROOT

Интерфейс реализован на базе классов ROOT GUI (классы семейства TG*), являющегося частью стандартной инфраструктуры ROOT для построения оконных приложений. Для формирования окна используются типовые элементы:

1. контейнеры и layout-менеджеры (TGFrame, TGVerticalFrame, TGHorizontalFrame, TGLayoutHints) для компоновки;

2. вкладки (TGTab) для логической группировки;
3. элементы управления вводом (TGTextEntry для чисел/строк, чекбоксы и кнопки типа TGCheckButton, TGTextButton);
4. текстовые подписи (TGLabel) для документирования назначения полей.

Заклучение

В ходе научно-исследовательской работы проведено исследование событий рождения двухмюонного резонанса с массой около 28 ГэВ, обнаруженного в данных эксперимента CMS (Run 1 и Run 2).

Основные результаты работы:

1. Исследована и применена последовательность критериев отбора для выделения сигнальных событий на фоне основных процессов Дрелла-Яна, $t\bar{t}$ и рождения пары бозонов.
2. Построено распределение инвариантной массы пары мюонов и выполнено его фитирование с использованием метода максимального правдоподобия. Сигнал моделировался сверткой Брейта-Вигнера и Гаусса, фон — полиномом.
3. С помощью F-теста обоснован выбор степени полинома для описания фона. Дополнительно проведена проверка устойчивости фитирующей функции с использованием игрушечных экспериментов; значимого систематического смещения положения сигнального пика не обнаружено.
4. Для области $SR2^{mod}$ выполнена оценка вклада событий с одним и двумя b -кварками. Получено, что основной вклад в отобранные события дает конфигурация с одним b -кварком и одной легкой струей.
5. Для данных Run 2 при $\sqrt{s} = 13$ ТэВ выполнена оптимизация отборов по поперечному импульсу второго мюона и бусту димюонной системы. Максимальная значимость получена при требованиях $p_T^{\mu 1} > 25$ ГэВ, $p_T^{\mu 2} > 18$ ГэВ и $p_T^{\mu\mu} / m_{\mu\mu} > 1,7$.
6. Проведен анализ распределения недостающего поперечного импульса (p_T^{miss}) в сигнальной области. Обнаружено, что в событиях с $m_{\mu\mu} \approx 28$ ГэВ наблюдается ненулевой p_T^{miss} , что свидетельствует о наличии в конечном состоянии частиц, не регистрируемых детектором.
7. Проведено сравнение распределения p_T^{miss} с феноменологической моделью рождения тяжелого состояния с последующим распадом в димюонный резонанс и невидимые частицы. Получено качественное согласие формы распределения модели с данными Run 1.
8. Реализован графический интерфейс пользователя для улучшения пользовательского опыта при многократном запуске скрипта с одинаковым функционалом, но разными параметрами отбора.

Таким образом, результаты исследования указывают на нетривиальный характер событий с инвариантной массой пары мюонов около 28 ГэВ, проявляющийся в ассоциации с b -струями и наличии недостающего поперечного импульса. В будущем планируется провести моделирование рождения пары мюонов для предложенной теоретической модели и сравнить результат с экспериментальными данными.

Список литературы

- [1] CMS Collaboration. Search for resonances in the mass spectrum of muon pairs produced in association with b quark jets in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV. *JHEP*, 11:161, 2018.
- [2] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [3] V. Khachatryan et al. The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV at the LHC. *JINST*, 8:P11002, 2013.
- [4] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [5] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018.