

Использование моделей максимального правдоподобия для измерения интегрального поперечного сечения ассоциированного рождения двух векторных бозонов в протон-протонных столкновениях на эксперименте АТЛАС

Петухов А.М.

Научный руководитель — к.ф.-м.н. Солдатов Е.Ю.

Москва, 2023

Введение

Взаимодействия векторных бозонов определяются кинематическим слагаемым лагранжиана Стандартной модели (СМ).

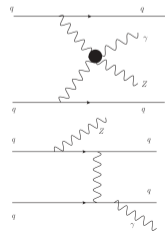
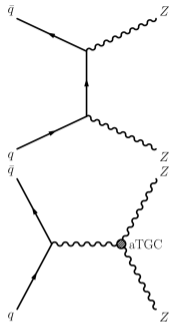
При нарушении калибровочной симметрии $SU(2)_L \times SU(1)_Y \rightarrow U(1)_{\text{ЭМ}}$ остаются доступными только взаимодействия

- ▶ W^+W^-Z и $W^+W^-\gamma$;
- ▶ $W^+W^-W^+W^-$, W^+W^-ZZ , $W^+W^-Z\gamma$, $W^+W^-\gamma\gamma$.

Взаимодействия трёх (четырёх) векторных бозонов изучаются в pp -столкновениях в процессах рождения VV ($VVjj$), $V = W/Z/\gamma$.

Изучение таких процессов позволяет

- ▶ проверить предсказания СМ;
- ▶ производить не прямые поиски эффектов, предсказываемых теориями, выходящими за пределы СМ («новой физики»).



Исследуемые процессы

В данной работе исследовались следующие процессы в pp -столкновения:

- ▶ ЭС-рождение $Z(\rightarrow \nu\bar{\nu})\gamma jj$, $\mathcal{O}(\alpha^4)$;
- ▶ $ZZ \rightarrow \ell^-\ell^+\nu\bar{\nu}$, где $\ell = e$ или μ

Мода распада $Z \rightarrow \nu\nu$ в сравнении с $Z \rightarrow \ell\ell$

- ▶ Обладает большей вероятностью (20.00 % и 6.72 %) \rightarrow более строгие ограничения на модели «новой физики».
- ▶ Сложнее отделяется от фоновых процессов из-за сложности измерения E_T^{miss} .

Существующие измерения

$Z(\rightarrow \nu\bar{\nu})\gamma jj$, АТЛАС, 139 fb^{-1} , $\sqrt{s} = 13 \text{ ТэВ}$. $E_T^\gamma \in [15; 110] \text{ ГэВ}$. Обнаружен, нет ограничений на новую физику.

$Z(\rightarrow \ell\ell)\gamma jj$, CMS, 137 fb^{-1} , $\sqrt{s} = 13 \text{ ТэВ}$. Обнаружен.

$ZZ \rightarrow \ell^-\ell^+\nu\bar{\nu}$, АТЛАС, 36.1 fb^{-1} , $\sqrt{s} = 13 \text{ ТэВ}$.

$ZZ \rightarrow \ell^-\ell^+\ell'^-\ell'^+$, АТЛАС, 139 fb^{-1} , $\sqrt{s} = 13 \text{ ТэВ}$ и CMS, 137 fb^{-1} , $\sqrt{s} = 13 \text{ ТэВ}$

Цели и задачи

Цель

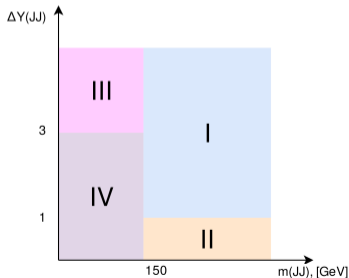
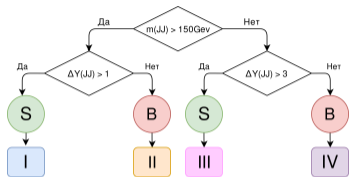
Используя данные pp -столкновений с $\sqrt{s} = 13$ ТэВ собранные детектором АТЛАС и соответствующие 139 fb^{-1} , измерить интегральное сечение процессов рождения ZZ и ЭС-рождения $Z\gamma jj$.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Разработать **метод разделения сигнальных и фоновых событий на основе алгоритма композиции деревьев решений** с целью повышения значимости измерения числа фоновых событий.
2. На основе **метода максимального правдоподобия** разработать метод учета статистических погрешностей и их взаимосвязи с целью измерения поперечного сечения.
3. Измерить **интегральное поперечное сечение** процесса электрослабого рождения $Z(\rightarrow \nu\bar{\nu})\gamma jj$ в протон-протонных столкновениях с энергией в СЦИ 13 ТэВ.
4. Измерить **силу сигнала** процесса рождения $ZZ \rightarrow ll\nu\nu$ в протон-протонных столкновениях с энергией в СЦИ 13 ТэВ на данных Азимова.

Композиция деревьев решений

Деревья решений



Бустинг

Пошаговое приближение $Pr(\text{сигнал}|\mathbf{x}_i)$ суммой откликов слабых классификаторов

- ▶ приоритет получают события, неправильно классифицируемые суммой предыдущих слагаемых;
- ▶ вклад слагаемого зависит от числа сделанных им ошибок и скорости обучения (параметр бустинга).

Преимущества композиции деревьев решений:

- ▶ данные не требуют предварительной обработки;
- ▶ хорошо справляется с избыточными переменными;
- ▶ мало параметров;

Общая идея:

Кинематические переменные системы частиц



Одна непрерывная переменная в отрезке $[-1; 1]$
(отклик алгоритма)

Метод максимального правдоподобия

Метод максимального правдоподобия используется для

- ▶ Оценки сечения рождения процессов.
- ▶ Оценки значимости полученного измерения.

Схема метода:

- ▶ Предсказанное число событий процесса p ν^p и измеренное число событий $N^{\text{данн.}}$
- ▶ Статистическая модель $\mathcal{L}(\nu) = \mathcal{L}(N^{\text{данн.}} | \nu)$
- ▶ Оценка полученного числа событий $\hat{\nu}_s = \arg[\max \mathcal{L}(\nu)]$

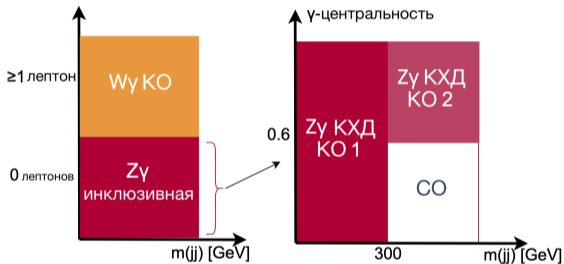
Для простоты используют **силу сигнала** $\mu = \nu_{\text{изм.}}^s / \nu_{\text{СМ}}^s = \sigma_{\text{дов., изм.}}^s / \sigma_{\text{дов., СМ}}^s$

Исследуемое фазовое пространство разбивается на непересекающиеся **сигнальные области** (СО) и **контрольные области** (КО), используемые для оценки μ и μ_b

Систематические погрешности моделируются при помощи доп. параметров θ , влияющих на отклик через $\eta(\theta)$ и ограниченных измерениями $\mathcal{L}(\theta)$.

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Конечное состояние и фоновые процессы

Конечное состояние: высокоэнергетичный фотон ($E_T^\gamma > 150$ ГэВ), большое значение недостающего поперечного импульса ($E_T^{\text{miss}} > 120$ ГэВ) и минимум 2 адронные струи с $p_T > 50$ ГэВ.



Центральность фотона

$$\zeta_\gamma = \left| \frac{y(\gamma) - \frac{y(j_1) + y(j_2)}{2}}{y(j_1) - y(j_2)} \right|$$

Сигнальный процесс

Z -рождение $Z\gamma jj$ [22.5%]

Фоны, оцениваемые

в контрольных областях

$KХД$ -рождение $Z\gamma jj$ [36%]

$W\gamma, tt\gamma$ [38%]

из данных

$e \rightarrow \gamma$ [6%]

$j \rightarrow \gamma$ [5.5%]

Неправильное измерение E_T^{miss} [2%]

напрямую из МК

$Z(\ell\ell) + \gamma$ [$< 1\%$]

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Классификатор машинного обучения

Используемые переменные:

1. m_{jj} ;
2. $\Delta Y(j_1, j_2)$;
3. E_T^{miss} ;
4. p_T -баланс;
5. $\eta(j_2)$;
6. $p_T(j_1)$;
7. $\eta(\gamma)$;
8. p_T -баланс (усеченный);
9. N_{jets} ;
10. $\sin\left(\left|\frac{\Delta\varphi(j_1, j_2)}{2}\right|\right)$;
11. $\Delta Y(j_1, \gamma)$.

Настройка классификатора

Переменные отобраны методом пошагового отбора «N-1» из списка в 26 переменных.

Достоверность моделирования переменных и их корреляций проверена в контрольных областях.

Параметры классификатора настроены с целью получения максимальной значимости при стабильности классификатора.

Сравнение фиксированного отбора по отклику с наилучшими отборами по m_{jj} , p_T -баланс, $N_{\text{струи}}$, $y(j_2)$

	S	B	$S/\sqrt{S+B}, \sigma$
Фикс. отборы	25.9 ± 0.1	38.2 ± 1.1	3.23 ± 0.03
Классификатор	33.3 ± 0.2	43.0 ± 1.5	3.81 ± 0.04

Использование классификатора позволяет увеличить оценку значимости на 18 %.

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Модель максимального правдоподобия

Нормировочные коэффициенты и погрешности

Процесс	Норм. коэфф.	Систематическая погрешность	
		Собственная	Общая
$Z_{\gamma jj}$ EWK	$\mu(Z_{\gamma EWK}), POI$	Теоретическая, интерференция	
$Z_{\gamma jj}$ QCD	$\mu(Z_{\gamma QCD})$	Теоретическая	
W_{γ} EWK		Теоретическая	Экспериментальные
W_{γ} QCD	$\mu(W_{\gamma})$	Теоретическая	
$t\bar{t}\gamma$		Теоретическая	
$Z(l\bar{l})+\gamma$		Теоретическая	
$e \rightarrow \gamma$		Плоская, из данных	
$j+\gamma$		Плоская, из данных	
$j-\gamma$		Плоская, из данных	

3 контрольные области
распределение по m_{jj}
1 сигнальная область
распределение по отклику
классификатора

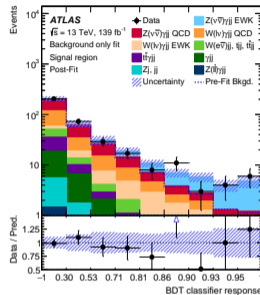
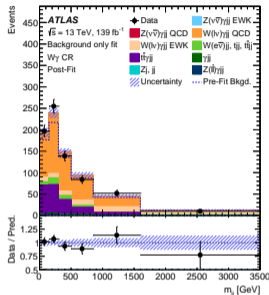
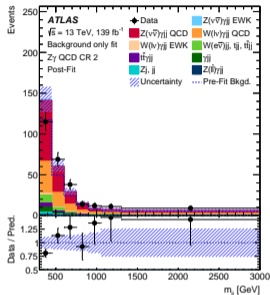
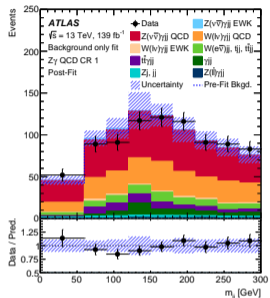
Дополнительные погрешности

- ▶ Интерференция между процессами ЭС- и КХД-рождения $Z\gamma jj$.
- ▶ Моделирование m_{jj} в сигнальном регионе.
- ▶ Для моделирования фона от «пайлапа»¹ ко всем остальным процессам применялась погрешность в 1.9 %.

¹ «пайлап» — pp -взаимодействия, сопровождающие основное во время одного пересечения пучков

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Результаты

Распределения после максимизации правдоподобия



Измеренная сила сигнала

$$\mu_{Z\gamma\text{ЭС}} = 0.78_{-0.23}^{+0.25} \text{ (стат.) }_{-0.17}^{+0.21} \text{ (сист.)}$$

Предсказанное сечение

$$\sigma_{Z\gamma\text{ЭС}}^{\text{предск.}} = 0.98 \pm 0.02 \text{ (стат.)} \pm 0.09 \text{ (шкал.)} \pm 0.02 \text{ (структ.) фб.}$$

Измеренное сечение

$$\sigma_{Z\gamma\text{EWK}} = 0.77_{-0.30}^{+0.34} \text{ фб} = 0.77_{-0.23}^{+0.25} \text{ (стат.) }_{-0.18}^{+0.22} \text{ (сист.) фб.}$$

С измеренной (ожидаемой) значимостью 3.2σ (3.7σ).

$ZZ \rightarrow \ell^- \ell^+ \nu \bar{\nu}$. Конечное состояние и фоновые процессы.

Конечное состояние: пара противоположно заряженных лептонов (e или μ) с $p_T > 20$ ГэВ и большое значение недостающего поперечного импульса ($E_T^{\text{miss}} > 70$ ГэВ).

Сигнальный процесс

ZZ [18%]

Фоны, оцениваемые

в контрольных областях

$Z(\ell\ell)$ +струи (неправильное измерение E_T^{miss}) [49%]

«Нерезонансное» (НР) рождение пары $\ell^- \ell^+$ [20%]

$WZ, Z \rightarrow \ell\ell$ [13%]

напрямую из МК

$VVV, 4\ell, \ell\ell qq$ [$< 1\%$]

Отбор	CO	WZ KO	НР KO	Z+струи KO
E_T^{miss} -значимость	> 7	> 7	> 7	[4; 7]
Аромат пары лептонов	$e^- e^+ / \mu^- \mu^+$	$e^- e^+ / \mu^- \mu^+$	$e^- \mu^+ / \mu^- e^+$	$e^- e^+ / \mu^- \mu^+$
$N_{\text{лептоны}}$	= 2	= 3	= 2	= 2
m_T^W , ГэВ	—	> 60	—	—

$ZZ \rightarrow \ell^- \ell^+ \nu \bar{\nu}$. Модель максимального правдоподобия

Процесс	Норм. коэфф.	Систематическая погрешность	
		Собственная	Общая
ZZ	$\mu(\text{ZZ}), \text{POI}$	Теоретическая	
Z+струи	$\mu(\text{Zj})$	Теоретическая	
HP рождение	$\mu(\text{HP})$	Теоретическая	Экспериментальные
WZ	$\mu(\text{WZ})$	Теоретическая	
Прочие		Теоретическая	

3 контрольные области

Распределение по p_{T}^Z

1 сигнальная область

Распределение по отклику классификатора

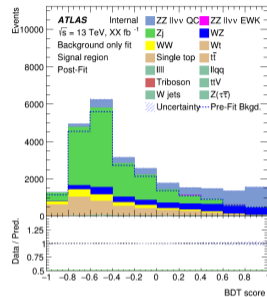
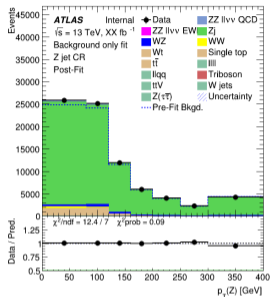
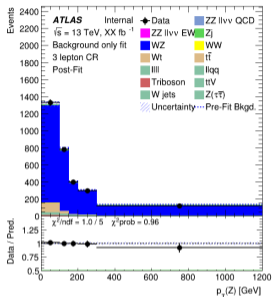
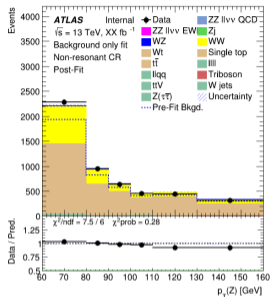
Поскольку анализ находится не на финальной стадии, наблюдаемые данные в сигнальной области не используются.

Оценка силы сигнала производится по следующей схеме:

1. Производится максимизация правдоподобия в КО, с целью оценить нормировочные коэффициенты фонов μ_b и доп. параметры сист. погрешностей θ .
2. Полученные значения μ_b и θ через функции откликов η применяются к предсказанным распределениям.
3. Полученные распределения используются во всех областях, вместо наблюдаемых, и заново происходит максимизация правдоподобия.

ZZ → ℓ⁻ℓ⁺νν̄. Результаты

Распределения после первого шага оценки



Коэффициент	Величина
-------------	----------

μ_{ZZ}	$1.00^{+0.03}_{-0.03}$ (стат.) $^{+0.06}_{-0.05}$ (сист.)
μ_{Zj}	$1.13^{+0.01}_{-0.01}$ (стат.) $^{+0.06}_{-0.06}$ (сист.)
μ_{NP}	$1.15^{+0.02}_{-0.02}$ (стат.) $^{+0.05}_{-0.05}$ (сист.)
μ_{WZ}	$0.97^{+0.02}_{-0.02}$ (стат.) $^{+0.06}_{-0.05}$ (сист.)

Ожидаемая значимость
26.1 σ

Заключение

В ходе данной работы

1. Был разработан метод разделения сигнальных и фоновых процессов на основе композиции деревьев решений, созданной при помощи бустинга, позволивший увеличить оценку значимости разделения на 18 %.
2. Был разработан метод оценки поперечного сечения на основе метода максимального правдоподобия, позволивший учесть вклад всех систематических и статистических погрешностей и их взаимосвязи.
3. Было измерено интегральное поперечное сечения процесса электрослабого рождения $Z(\rightarrow \nu\bar{\nu})\gamma jj$ в протон-протонных столкновениях с энергией в СЦИ 13 ТэВ, которое составило
$$\sigma_{Z\gamma EWK} = 0.77^{+0.25}_{-0.23} \text{ (стат.) }^{+0.22}_{-0.18} \text{ (сист.) фб.}$$
4. На данных Азимова была получена сила сигнала процесса рождения $ZZ \rightarrow ll\nu\nu$ в протон-протонных столкновениях с энергией в СЦИ 13 ТэВ, которая составила
$$\mu_{ZZ} = 1.00^{+0.03}_{-0.03} \text{ (стат.) }^{+0.06}_{-0.05} \text{ (сист.)}.$$

Спасибо за внимание!

Ответы на замечания рецензентов. Храмов Е.В. 1/2

Во введении написано, что выбор именно процессов рождений ZZ и $Z\gamma$ для исследования позволяет «получить наиболее строгие ограничения на параметры моделей, предсказывающих новые взаимодействия электронейтральных векторных бозонов». Неясно, рождения каких других пар могут дать возможность получить менее строгие ограничения на взаимодействия электронейтральных векторных бозонов?

Для 4 векторных бозона: $ZZjj$.

Для 3 векторных бозонов конечное состояние $Z\gamma$ позволяет получить более строгие ограничения на указанные взаимодействия. Однако комбинация результатов с результатами, полученными для ZZ позволит их усилить.

Ответы на замечания рецензентов. Храмов Е.В. 2/2

Из главы 4 остаётся непонятным, какая систематическая погрешность является доминирующей в исследовании процесса рождения ZZ

Поскольку для процесса рождения ZZ значения погрешностей брались из работы [JHEP 10 \(2019\) 127](#) и были одинаковыми для каждого процесса, то наибольший вклад в погрешность силы сигнала вносит погрешность, наиболее влияющая на число событий — погрешность на структурные функции протона.

Теоретические		Экспериментальные	
Структ. функц.	3.5 %	Лептоны	2.0 %
Масштабы μ_R и μ_F	2.0 %	Струи	2.0 %
Моделирование	2.0 %	E_T^{miss}	1.1 %

Ответы на замечания рецензентов. Поликарпов С.М.

Термин «данные Азимова» крайне редко встречается в русскоязычной литературе и должен быть подробно определён. В работе Петухова А. М. данные Азимова определяются как «данные, при использовании на которых метода максимального правдоподобия, всегда будут получены истинные значения оцениваемых параметров». Из этого непонятно, как такие данные получены, и что подразумевается под «истинными значениями параметров»

Данные Азимова можно получить, сложив для каждого рассматриваемого бина все значения предсказанных чисел событий с влиянием на них нормировочных коэффициентов и параметров систематических погрешностей, .

$$\bar{\mu}\nu_i^s\eta^s(\bar{\theta}) + \nu_i^b\eta^b(\bar{\theta})$$

В таком случае

$$\hat{\mu}, \hat{\theta} = \arg[\max \mathcal{L}(\mu, \theta)] = \bar{\mu}, \bar{\theta}$$

Для $\bar{\theta}$ могут использоваться значения

- ▶ по-умолчанию, т.е. никак не влияющие на предсказанное число событий. Например, $\mu_b = 1$;
- ▶ полученные при максимизации правдоподобия в особых условиях. Например, только в контрольных областях.

Ответы на замечания рецензентов. Поликарпов С.М.

Не обсуждается систематическая погрешность, связанная с выбором разбиения переменных на бины.

Данная погрешность не учитывалась в работе.

Путём варьирования числа бинов в контрольных регионах для исследования $Z_{\gamma jj}$, было получено, что их влияние на погрешность измеренного сечения не превышает 7 %.

Этот эффект меньше основных вкладов в погрешность — теоретических погрешностей на сигнальные и фоновые процессы.

Ответы на замечания рецензентов. Поликарпов С.М.

Не все критерии отбора и разбивка на контрольные / сигнальные области достаточно аргументированы

Общие критерии для отбора событий:

- ▶ Соответствие геометрическим параметрам детектора, порогам регистрации частиц и используемым триггерам.
- ▶ Уменьшение числа фоновых событий.

Контрольные области отбираются так, чтобы

- ▶ Получившиеся области фазового пространства не пересекались друг с другом и с СО. Это делается для упрощения использования метода максимального правдоподобия.
- ▶ Вклад рассматриваемого фонового процесса был максимальным.

Научная новизна

1. Впервые использована комбинация метода разделения сигнальных и фоновых событий на основе алгоритма композиции деревьев решений и максимального правдоподобия для измерения интегральных поперечных сечений процессов ассоциированного электрослабого рождения $Z(\rightarrow \nu\bar{\nu})\gamma jj$ и $ZZ \rightarrow \ell\ell\nu\nu$ в протон-протонных столкновениях.
2. Впервые разработаны специальные методы оценки погрешностей, типичных только для исследуемого конечного состояния процесса электрослабого рождения $Z\gamma jj$.
3. Впервые измерено со значимостью больше 3σ сечение процесса ЭС-рождения $Z(\rightarrow \nu\bar{\nu})\gamma jj$ в области с большими значениями поперечного импульса фотона, $E_T^\gamma > 150$ ГэВ.

Практическая значимость

1. Разработанные методы разделения сигнальных и фоновых процессов и измерения интегрального поперечного сечения могут быть использованы в исследованиях любых других многобозонных конечных состояний в эксперименте АТЛАС.
2. Разработанный метод измерения интегрального поперечного сечения может быть использован как отправная точка для измерения дифференциальных поперечных сечений, а также для постановки ограничений на параметры теорий, выходящих за пределы СМ.

Апробация работы

Основные результаты работы докладывались на:

1. Конференции «41st International Conference on High Energy Physics», Болонья, Италия, 2022 г.
2. Конференции «10th International Conference on New Frontiers in Physics», Колумбари, Греция, 2021 г.
3. Московской международной школе физики, Московская область, Россия, 2019 г.
4. Конференции «The 3rd International Conference on Particle Physics and Astrophysics», Москва, Россия, 2019 г.

Публикации по теме НКР. 1/2

- A1. Measurement of Electroweak $Z(\nu\bar{\nu})\gamma jj$ Production and Limits on Anomalous Quartic Gauge Couplings in Pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector [Текст] / G. Aad [et al.] // Journal of High Energy Physics. — 2023. — June. — Vol. 2023, no. 6. — P. 82. — (Visited on 06/18/2023).
- A2. *Petukhov A.* Measurements of Quartic Coupling and Vector Boson Scattering in ATLAS [Текст] // Proceedings of 41st International Conference on High Energy Physics — PoS(ICHEP2022). — Bologna, Italy : Sissa Medialab, 06/2023. — P. 1068. — (Visited on 03/01/2023).
- A3. *Belyaev N. L., Petukhov A. M., Soldatov E. Y.* The Automation of Choosing of the Optimal Kinematic Variables for Discrimination of the Electroweak $Z\gamma$ Production [Текст] // Journal of Physics: Conference Series. — 2019. — Nov. — Vol. 1390, no. 1. — P. 012040. — (Visited on 03/01/2023).

- A4. *Belyaev N., Petukhov A., Soldatov E.* New Methods of Distinguishing the Associated $Z\gamma$ Production [Текст] // EPJ Web of Conferences. — 2019. — Vol. 222. — P. 02015. — (Visited on 03/01/2023).
- A5. *Petukhov A. M., Soldatov E. Y.* Usage of Machine Learning for the Separation of Electroweak and Strong $Z\gamma$ Production at the LHC Experiments [Текст] // Journal of Physics: Conference Series. — 2017. — Dec. — Vol. 934, no. 1. — P. 012028. — (Visited on 03/01/2023).

Взаимодействия векторных бозонов

$$\mathcal{L}_{\text{ЭС, кин.}} = -\frac{1}{4} \mathbf{G}_{\mu\nu} \cdot \mathbf{G}^{\mu\nu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu},$$

где $G_{\mu\nu}^a$ и $F_{\mu\nu}$ — напряженности полей \mathbf{W}_μ и A_μ , определяющиеся как

$$G_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - ig\varepsilon_{bc}^a W_\mu^b W_\nu^c, \quad \text{где } a, b, c = 1, 2, 3,$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

$$W^1 = \frac{W^+ + W^-}{\sqrt{2}},$$

$$W^2 = \frac{W^+ - W^-}{\sqrt{2}i},$$

$$W^3 = Z \cos \theta_w + \gamma \sin \theta_w.$$

$$ig\varepsilon_{bc}^a \partial_\mu W_\nu^a W_\mu^b W_\nu^c \rightarrow W^i W^j W^k, \quad \text{где } i, j, k = 1, 2, 3; i \neq j, j \neq k, i \neq k,$$

$$(ig\varepsilon_{bc}^a W_\mu^b W_\nu^c)^2 \rightarrow W^i W^j W^i W^j, \quad \text{где } i, j = 1, 2, 3; i \neq j.$$

- ▶ $W^+ W^- Z$ и $W^+ W^- \gamma$;
- ▶ $W^+ W^- W^+ W^-$, $W^+ W^- Z Z$, $W^+ W^- Z \gamma$, $W^+ W^- \gamma \gamma$.

Метод максимального правдоподобия

Полная формула

$$\mathcal{L}(\mu, \theta) = \prod_r^{\text{обл.}} \left[\prod_i^{\text{бины} \in r} \text{Pois}(N_i^{\text{данн.}} | \mu \nu_i^s \eta^s(\theta) + \nu_i^b \eta^b(\theta)) \right] \cdot \prod_i^{\text{доп. парам.}} \mathcal{L}(\theta_i),$$

где

- ▶ $N^{\text{данн.}}$ и ν — число наблюдаемых и предсказанных событий.
- ▶ μ — сила сигнала $\mu = \nu_{\text{изм.}}^s / \nu_{\text{СМ}}^s = \sigma_{\text{дов., изм.}}^s / \sigma_{\text{дов., СМ}}^s$
- ▶ θ — норм. коэфф. фона и параметры систематических погрешностей.
- ▶ $\eta(\theta)$ — функции отклика θ , описывающие их влияние на число событий.
- ▶ $\mathcal{L}(\theta_i)$ — измерение, накладывающее ограничение на дополнительный параметр θ_i .

Вычисление значимости

$$q(\mu, \hat{\mu}, \hat{\theta}) = -2 \ln \lambda(\mu, \hat{\mu}, \hat{\theta}) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \quad (1)$$

где

- ▶ $\lambda(\mu, \hat{\mu}, \hat{\theta})$ — отношение правдоподобий.
- ▶ $\hat{\theta}(\mu)$ — набор значений дополнительных параметров, минимизирующий $-\ln \mathcal{L}(\mu, \theta)$ для данного значения μ .

Значимость измерения $\hat{\mu}$ в экспериментальных данных и оценки медианной ожидаемой значимости могут быть вычислены как

$$\begin{aligned} Z_{\text{набл.}}^{\text{изм.}} &= \sqrt{q(\mu = 0)}, \\ Z_{\text{набл.}}^{\text{ожид.}} &= \sqrt{q(\mu = 1)_A}, \end{aligned} \quad (2)$$

где $q(\mu = 1)_A$ вычисляется на данных Азимова.

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Отбор событий

Адронные струи

$N_{\text{jets}} \geq 2$;
 $p_T^{\text{jet}} > 50$ ГэВ.

Фотон

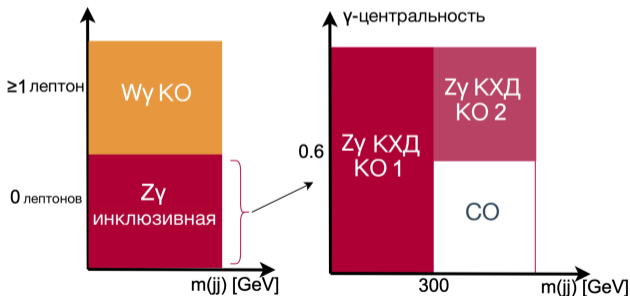
$N_\gamma = 1$;
 $E_T^\gamma > 150$ ГэВ;
 $E_T^{\text{cone40}} < 2.45$ ГэВ + $0.022 \cdot E_T^\gamma$;
 $p_T^{\text{cone20}} < 0.05 \cdot E_T^\gamma$;
 $|z(\gamma)| < 250$ мм;
 Высокое качество идентификации.

Недостающий поперечный импульс

$\vec{p}_T^{\text{miss}} = -(\sum_{\text{obj}} \vec{p}_T^{\text{obj}} + \vec{p}_T^{\text{SoftTerm}})$;
 $|\vec{p}_T^{\text{miss}}| = E_T^{\text{miss}} > 120$ ГэВ;
 Значимость измерения \vec{p}_T^{miss} ,
 $S > 12$;
 $|\vec{p}_T^{\text{SoftTerm}}| < 16$ ГэВ.

Дополнительные ограничения

Вето на e^\mp и μ^\mp в событии;
 $\Delta\varphi(\vec{p}_T^{\text{miss}}, \gamma) > 0.4$;
 $\Delta\varphi(\vec{p}_T^{\text{miss}}, j_1) > 0.3$;
 $\Delta\varphi(\vec{p}_T^{\text{miss}}, j_2) > 0.3$.



$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Число событий в рассматриваемых областях

После максимизации правдоподобия

	$W\gamma$ CR	$Z\gamma$ QCD CR 1	$Z\gamma$ QCD CR 2	Signal region
$Z(\nu\bar{\nu})\gamma jj$ EWK	0.108 ± 0.028	11.0 ± 4.3	4.0 ± 2.2	37 ± 14
$Z(\nu\bar{\nu})\gamma jj$ QCD	1.04 ± 0.46	394 ± 84	143 ± 32	133 ± 39
$W(l\nu)\gamma jj$ QCD	425 ± 63	237 ± 71	76 ± 24	91 ± 30
$W(l\nu)\gamma jj$ EWK	63 ± 12	14.3 ± 2.7	4.5 ± 1.2	24.6 ± 4.9
$W(e\nu)jj, tjj, t\bar{t}jj$	39.8 ± 2.5	70.1 ± 4.1	17.9 ± 1.3	22.5 ± 1.5
$t\bar{t}\gamma jj$	193 ± 57	57 ± 20	9.1 ± 3.4	21.3 ± 7.6
γjj	4.8 ± 7.4	52 ± 36	8 ± 11	20 ± 17
Zj, jj	0.06 ± 0.66	20 ± 14	5.9 ± 6.9	6.6 ± 7.8
$Z(l\bar{l})\gamma jj$	8.6 ± 2.5	6.8 ± 2.0	2.04 ± 0.95	2.2 ± 1.3
Total	735 ± 30	863 ± 54	271 ± 25	357 ± 30
Data	737	849	268	356

$Z(\rightarrow \nu\bar{\nu})_{\gamma jj}$. BDT классификатор

Gradient Boosted Decision Trees

Parameter	Value
Number of trees	600
Shrinkage	0.05
Max tree depth	3
Number of cuts	20
Min node size	5%

Cross training

Datasets are split by event number into **odd** and **even**

	Train	Test & application
Odd algorithm	Odd events	Even events
Even algorithm	Even events	Odd events

Final response distribution is obtained by summing the response of even and odd events.

Variable selection

Selected variables:

- m_{jj}
- $\Delta Y(j_1, j_2)$
- E_T^{miss}
- p_T -balance
- $\eta(j_2)$
- $p_T(j_1)$
- $\eta(\gamma)$
- p_T -balance (reduced)
- N_{jets}
- $\sin(\Delta\varphi(j_1, j_2)/2)$
- $\Delta Y(j_1, \gamma)$

Considered variables:

- > $\eta(j_1)\eta(j_2)$
- > $\Delta R(j_1, j_2)$
- > $\Delta R(j_1, \gamma)$
- > $\Delta p_T(j_1, j_2)$
- > ζ_{jets}
- > $\Delta R(j_1, p_T^{\text{miss}})$
- > $\varphi(j_1)$
- > $\eta(j_1)$
- > $p_T(j_2)$
- > $\varphi(j_2)$
- > $p_T(\gamma)$
- > $\varphi(\gamma)$
- > $\varphi(p_T^{\text{miss}})$
- > $p_T(E_T^{\text{miss}}, \gamma)$

Variables were selected with "N-1" and "N+1" algorithms and checks for redundancy with correlation coefficients.

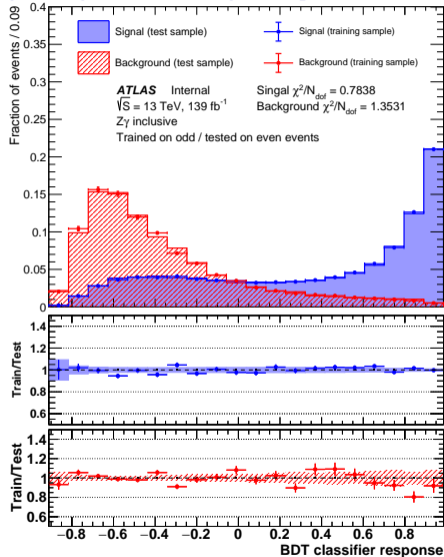
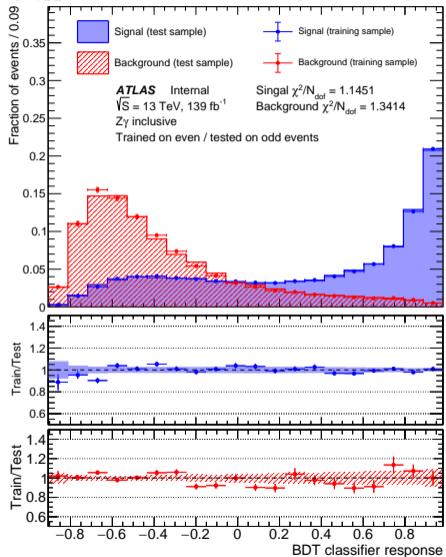
γ -centrality was not considered because it is used to define Z_γ QCD CR 2 and SR.

$$p_T\text{-balance} = \frac{|\bar{p}_T^{\text{miss}} + \bar{p}_T^\gamma + \bar{p}_T^{j_1} + \bar{p}_T^{j_2}|}{E_T^{\text{miss}} + E_T^\gamma + p_T^{j_1} + p_T^{j_2}}$$

$$p_T\text{-balance(reduced)} = \frac{|\bar{p}_T^\gamma + \bar{p}_T^{j_1} + \bar{p}_T^{j_2}|}{E_T^\gamma + p_T^{j_1} + p_T^{j_2}}$$

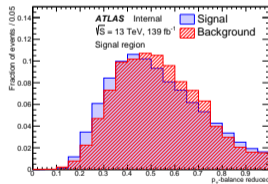
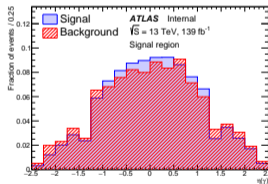
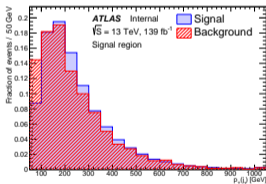
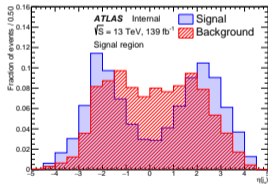
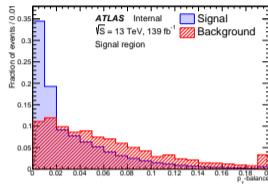
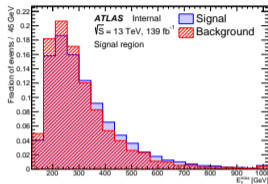
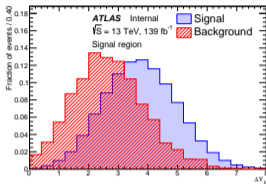
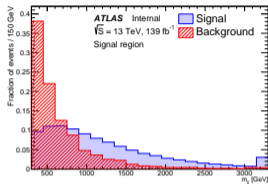
$$X\text{-centrality: } \zeta_X = \left| \frac{y(X) - \frac{y(\hat{a}_1) + y(\hat{a}_2)}{2}}{y(\hat{a}_1) - y(\hat{a}_2)} \right|$$

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. BDT классификатор. Проверка на переобучение



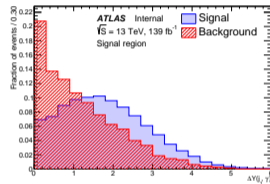
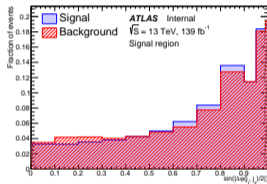
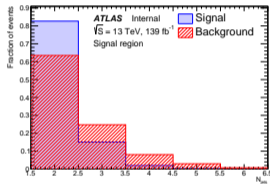
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Распределения по разделяющим переменным

1/2



$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Распределения по разделяющим переменным

2/2



$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Экспериментальные систематические погрешности

Type	Set
JES	30 NPs, <u>CategoryReduction</u>
JER	8 NPs, SimpleJER
MET	3 NPs
e/ γ	2 NP, <u>"1NP set"</u>
Muon	10 NPs
e efficiency	3 NPs
γ efficiency	2 NPs
Trigger efficiency	1 NP
JVT efficiency	1 NP
PRW	1 NP
Luminosity	1 NP
Pile-up bkg	1 NP

1 NP per systematic type for every background **estimated from the MC**, e.g. $\alpha(\text{JET_EffectiveNP_Detector1})$

The systematics are pruned using the following rules.

- The normalization part is dropped if the total effect on the event yield is $< 1\%$
- The shape part is dropped if there's no bins with effect on the event yield is $> 1\%$

See backup for more info

Flat 1.9% systematic to account for pile-up background event yield

Every background with **data-driven estimations** also have an NP for their estimated systematic (see slide 5) which results in 3 more NPs

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Теоретические систематические погрешности

	Modelling		Scale	Alt. PDF set	NNPDF + α_s
	Underlying Event + Parton showering	Sherpa vs MG			
Z γ jj EWK	✓		✓	✓	✓
Z γ jj QCD		✓	✓	✓	✓
W γ EWK	✓		✓	✓	✓
W γ QCD		✓	✓	✓	✓
t $\bar{t}\gamma$	✓		✓		✓
Z(l \bar{l})+ γ			✓	✓	✓

1 NP per combination of sample and systematic, e. g. α (tty UE+PS), α (W γ QCD alt PDF set), except for **scale**.

Scale:

4 NPs per sample for:

- W γ CR
- Z γ QCD CR1
- Z γ QCD CR2
- SR

Modelling

W γ QCD - no alternative sample available. Made with the Z γ QCD modelling systematic relative uncertainties (see backup for more info)

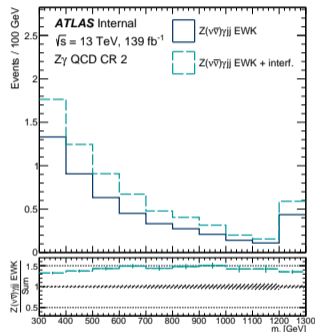
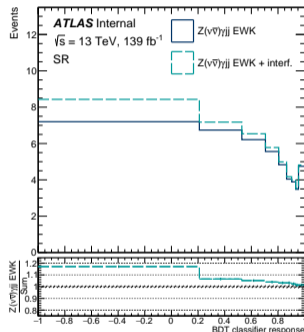
t $\bar{t}\gamma$ - only fastsim alternative sample available. Made by comparing fastsim nominal sample with fastsim alternative sample

$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Нетривиальные источники погрешностей

Пример: интерференция между ЭС- и КХД-рождением $Z(\rightarrow \nu\bar{\nu})\gamma jj$.

Оценивается путем создания отдельного Монте-Карло набора, соответствующего интерференционному слагаемому.

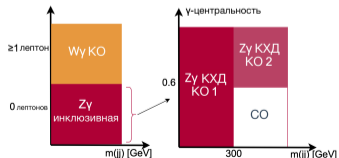
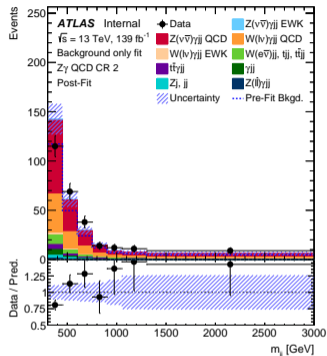
Variation	Onesided	Twosided	Symmetrical
+ σ	ZgEWK + interference	ZgEWK + interference	ZgEWK + interference
Nominal	ZgEWK	ZgEWK	ZgEWK
- σ	ZgEWK	ZgEWK + interference	ZgEWK - interference



Выбор симметричного подхода к моделированию создает более консервативную оценку погрешности, но позволяет избежать существования нескольких максимумов у функции правдоподобия.

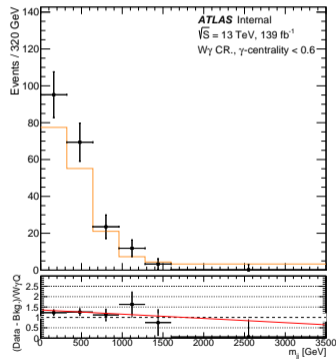
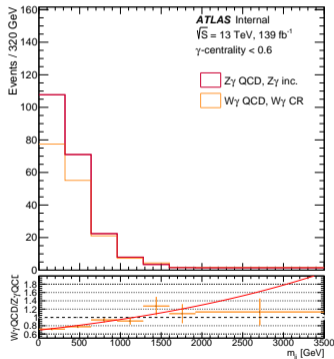
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Несоответствие в распределении по m_{jj}

$Z\gamma$ QCD CR 2



Процесс	% пред. событий
$Z\gamma$, QCD	45.41
$W\gamma$, QCD	30.74
$W(e\nu)$, top, $t\bar{t}$	7.88

Расхождение между предсказанием и данными в контрольной области $Z\gamma$ 2 может свидетельствовать об аналогичном расхождении в сигнальной области.

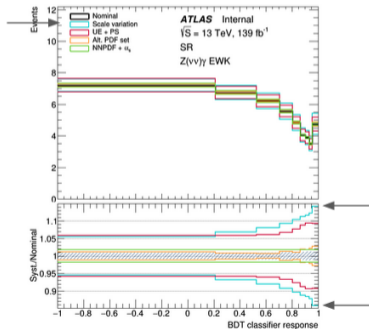


Результат: КХД-рождение $Z\gamma jj$ в SR смоделировано правильно с точностью до двух дополнительных погрешностей.

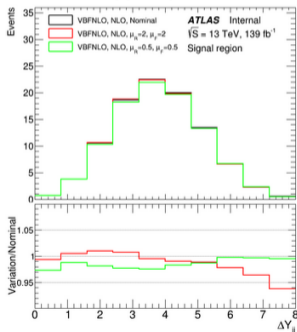
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Учет высоких порядков

Учёт петлевых КХД-поправок для сигнального процесса ЭС-рождения $Z\gamma jj$ позволяет значительно сократить погрешность связанную с приближением суммы более высоких порядков, не используемых в вычислении напрямую.

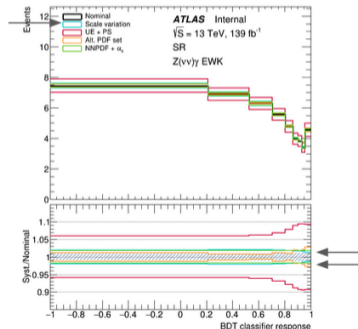
Без КХД-поправок



Погрешности в наборе с КХД-поправками

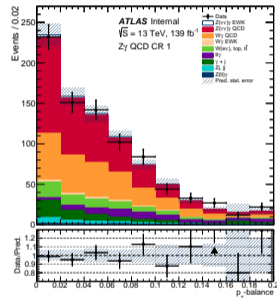
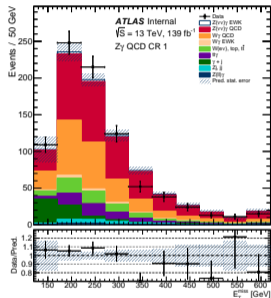
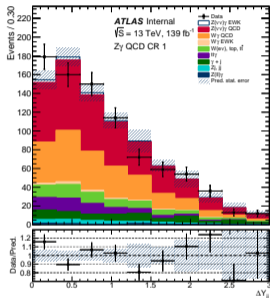
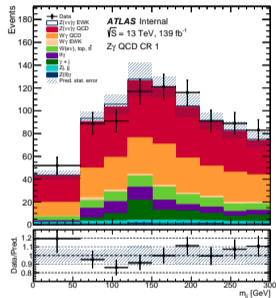


С учетом КХД-поправок



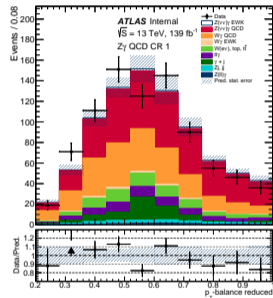
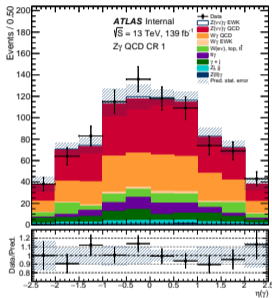
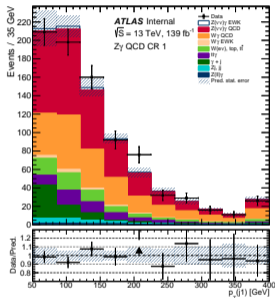
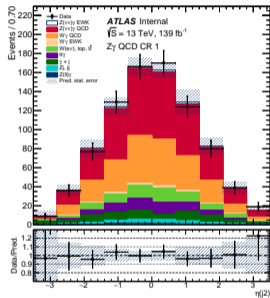
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Разделяющие переменные. Данные и моделирование

1/3



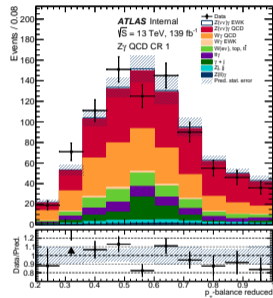
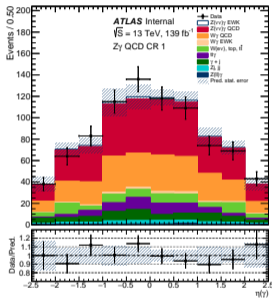
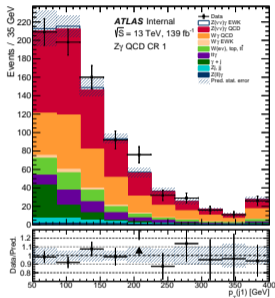
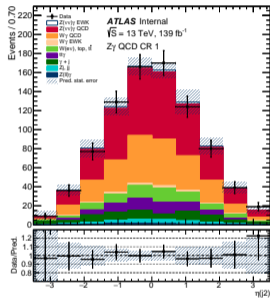
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Разделяющие переменные. Данные и моделирование

2/3



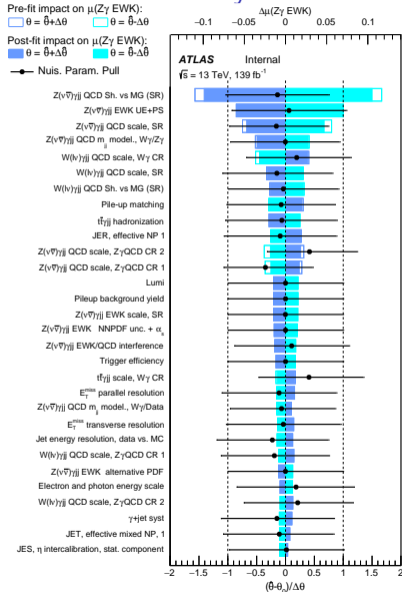
$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Разделяющие переменные. Данные и моделирование

3/3



$Z(\rightarrow \nu\bar{\nu})\gamma jj$. Вклад систематических погрешностей в ошибку сечения

Источник погрешности	$\Delta\sigma/\sigma$ [%]
Экспериментальные	
Струи	-3.2 / +3.4
Электроны и фотоны	-0.3 / +1.7
Мюоны	-0.4 / +0.5
E_T^{miss}	-1.8 / +2.2
Моделирование пайлапа	-1.7 / +3.2
Эффективность триггеров	-0.9 / +2.1
Светимость	-1.2 / +2.6
Теоретические	
$Z(\nu\bar{\nu})\gamma jj$ ЭС/КХД интерференция	-0.6 / +2.6
ЭС-рождение $Z(\nu\bar{\nu})\gamma jj$	-6 / +12
КХД-рождение $Z(\nu\bar{\nu})\gamma jj$	-15 / +16
Прочие процессы	-5.3 / +7.7
Прочие	
Оцененные из данных фоны	-0.9 / +1.2
Пайлап фон	-1.2 / +2.6
Моделирование m_{jj} для КХД $Z(\nu\bar{\nu})\gamma jj$	-4.4 / +4.4



ZZ. Отбор событий

Electrons

- ▶ Likelihood medium
- ▶ p_T lead > 30 ГэВ
- ▶ p_T sublead > 20 ГэВ
- ▶ $|\eta|$ calo cluster < 2.47
- ▶ $|\Delta(z_0) \cdot \sin(\theta)| < 0.5$ мм
- ▶ $|d_0\text{-significance}| < 5$
- ▶ Isolation WP [FixedCutLoose](#)
- ▶ Crack region veto
- ▶ Исключение пересечений с мюонами и струями

Muons

- ▶ Medium
- ▶ $|\eta| < 2.5$
- ▶ p_T lead > 30 ГэВ
- ▶ p_T sublead > 20 ГэВ
- ▶ Combined muons
- ▶ $|\Delta(z_0) \cdot \sin(\theta)| < 0.5$ мм
- ▶ $|d_0\text{-significance}| < 3$
- ▶ Isolation WP [PflowLoose_FixedRad](#)
- ▶ Исключение пересечений со струями

Jets

- ▶ AntiKt4EMPFLOW
- ▶ $p_T > 30$ ГэВ
- ▶ $|\eta| < 4.5$
- ▶ JVT > 0.5
- ▶ Event-level cleaning for LooseBad jets

E_T^{miss}

- ▶ Tight WP, rebuilt with METMaker using selected leptons and all calibrated jets

ZZ. Определение областей фазового пространства

Отбор	SR	WZ	NR	Z+струи
E_T^{miss} , ГэВ	> 70	—	> 70	> 70
$\Delta R(\ell\ell)$	< 2.2	—	< 2.2	< 2.2
$\Delta\varphi(\vec{p}_T^{\text{miss}}, \vec{p}_T^{\ell\ell})$	> 1.3	—	> 1.3	> 1.3
E_T^{miss} -значимость	> 7	> 7	> 7	[4; 7]
Аромат пары лептонов	$e^-e^+/\mu^-\mu^+$	$e^-e^+/\mu^-\mu^+$	$e^-\mu^+/\mu^-e^+$	$e^-e^+/\mu^-\mu^+$
$N_{\text{лептоны}}$	$= 2$	$= 3$	$= 2$	$= 2$
m_T^W , ГэВ	—	> 60	—	—

ZZ. BDT классификатор

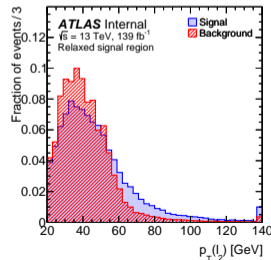
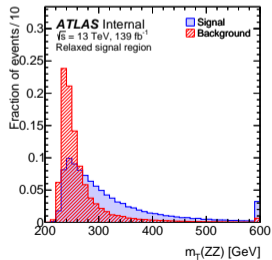
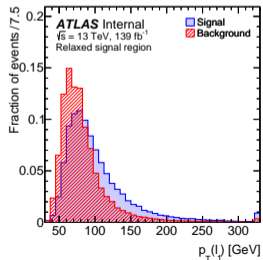
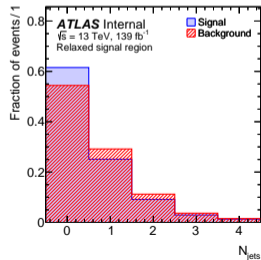
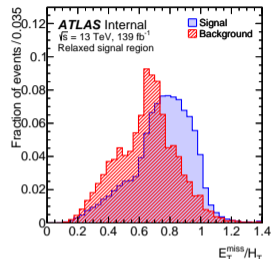
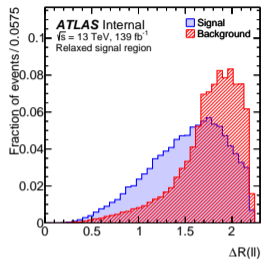
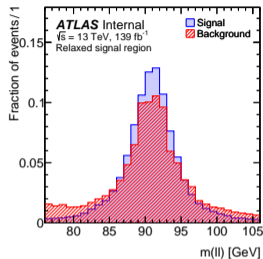
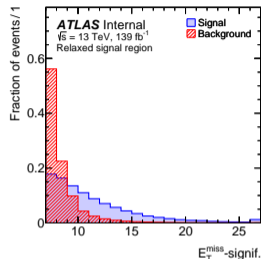
Рассматриваемые переменные

- ▶ E_T^{miss} ;
- ▶ E_T^{miss} -значимость;
- ▶ $m_T(ZZ)$;
- ▶ $\Delta R(\ell\ell)$;
- ▶ $p_T^{\ell 1}$;
- ▶ $p_T^{\ell 2}$;
- ▶ $p_T^Z = p_T^{\ell\ell}$
- ▶ $\frac{p_T^Z}{m_T(ZZ)}$;
- ▶ $H_T = \sum_i p_T^{\ell_i} + \sum_i p_T(j_i)$;
- ▶ E_T^{miss}/H_T ;
- ▶ $\rho_Z = \frac{p_T^Z}{p_T^{\ell 1} + p_T^{\ell 2}}$;
- ▶ $\Delta\varphi(\vec{p}_T^{\ell 1}, \vec{p}_T^{\ell 2})$;
- ▶ $\Delta\varphi(\vec{p}_T^{\text{miss}}, \vec{p}_T^{\ell\ell})$;
- ▶ $\frac{p_T^{\ell 1}}{p_T^{\ell 2}}$;
- ▶ $\Delta\eta(\ell\ell)$;
- ▶ $m(\ell\ell)$;
- ▶ y_Z ;
- ▶ N_{jets} ;
- ▶ $\frac{E_T^{\text{miss}} + \sum_i p_T(j_i) - p_T^Z}{p_T^Z}$;
- ▶ $V_T = \left| \sum_i \vec{p}_T^{\ell_i} + \sum_i \vec{p}_T^{\text{струи}_i} \right|$

Параметр	Значение
Скорость обучения	0.3
Макс. размер дерева	2
Мин. доля событий в листе	0.2 %
Число деревьев в комбинации	200

ZZ. Распределения по разделяющим переменным

1/2



ZZ. Распределения по разделяющим переменным

1/2

