О Б З О Р Ы ФИЗИКА АТОМНОГО ЯДРА И ЭЛЕМЕНТАРНЫХ ЧАСТИЦ

Принципы работы алгоритма реконструкции струй с использованием калориметрических и трековых детекторов при соударении частиц высокой энергии

О. Л. Кодолова^{1,*a*}, С. В. Образцов^{1,*b*}, И. Н. Варданян^{1,*c*}, А. Н. Никитенко², Н. В. Личковская², В. Б. Гаврилов², М. Д. Томс², Коллаборация СМS

Московский государственный университет имени М.В. Ломоносова, Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына. Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2.

²Институт теоретической и экспериментальной физики имени А.И. Алиханова,

НИЦ «Курчатовский институт» — ИТЭФ. Россия, 117218, Москва, ул. Б. Черемушкинская, д. 25.

E-mail: ^{*a*} olga.kodolova@cern.ch, ^{*b*} stepan.obraztsov@cern.ch, ^{*c*} cirina.vardanyan@cern.ch

Статья поступила 10.05.2017, подписана в печать 21.06.2017

На основе данных протон-протонных соударений при энергии и интегральной светимости в 19.8 fb⁻¹ было оценено разрешение по поперечному импульсу и неопределенность энергетической шкалы для струй, реконструированных с помощью алгоритма Jet-Plus-Track. Также была проведена оценка эффективности работы алгоритма идентификации струй из сигнального события и разделения струй на кварковые и глюонные.

Ключевые слова: рекострукция струй, калибровка струй, физика высоких энергий. УДК: 539.1.07. РАСS: 13.85.-t.

Введение

Алгоритм Jet-Plus-Track [1–3] корректирует энергию и направление адронных струй, реконструированных по энерговыделению в калориметрических ячейках детектора CMS [4] с использованием импульсов заряженных частиц, измеренных в трековом детекторе. Краткое описание принципа работы алгоритма Jet-Plus-Track приводится в разд. 1.

Для работы с алгоритмом Jet-Plus-Track требуется применение коррекций энергетической шкалы. Это связано с ограничением, связанным с неполным покрытием трековым детектором диапазона значений по псевдобыстроте, а также с нелинейностью отклика калориметрической системы детектора. Также необходимо вводить дополнительные коррекции, связанные с различием между реальными данными и результатами Монте-Карло-моделирования. Процедура коррекции приводится в разд. 3.

В период работы ускорителя Большой адронный коллайдер (БАК) при энергии на каждое пересечение пучков в среднем приходилось 23 дополнительных протон-протонных события, кроме того события, которое принято называть сигнальным. Типичное значение плотности дополнительных событий по поперечному импульсу составляет приблизительно 0.7 ГэВ на единицу площади (в плоскости $\eta-\varphi$). Следовательно, для каждой струи с параметром радиуса R = 0.5 [5] вклад дополнительных событий в конус струи составляет около 10 ГэВ. Кроме того, частицы от дополнительных событий могут

2 ВМУ. Физика. Астрономия. № 3

формировать достаточно жесткую струю, которая удовлетворяет кинематическим отборам, применяемым в физическом анализе. Выделение струй, относящихся к сигнальному событию, является важной методологической задачей. Решение этой задачи осуществляется путем анализа кинематических характеристик струи и ее составных частей, а также анализа формы струи с применением методов машинного обучения. Также данные характеристики могут быть использованы для разделения струй, инициированных глюоном, от струй, инициированных легким кварком. Описание кварк-глюонного дискриминатора приводится в разд. 6.

1. Описание алгоритма Jet-Plus-Track

Jet-plus-track-струи — это калориметрические струи с улучшенным энергетическим откликом и разрешением за счет использования информации от трекового детектора. Калориметрические струи реконструируются на основе энерговыделения в калориметрических ячейках, которые группируются в кластеры с помощью алгоритма anti- k_T [6] с радиусом R = 0.5. Далее треки заряженных частиц ассоциируются со струями по параметру пространственного разделения (в плоскости η - φ) между осью струи и вектором импульса заряженных частиц. Только треки, идущие от сигнального события, используются для ассоциирования со струей, найденной в калориметре. Сигнальное событие выбирается как событие с максимальным $\sum_{i \in vertex} p_{T,i}^2$, где $p_{T,i}$ — поперечный импульс трека. Принадлежность трека к событию определяется путем экстраполяции траектории трека к вершине, то есть к точке, в которой произошло изначальное соударение протонов. Трек считается ассоциированным со струей, если расстояние $R = \sqrt{(\eta_{\text{trk}} - \eta_{\text{jet}})^2 + (\varphi_{\text{trk}} - \varphi_{\text{jet}})^2}$ между треком и струей в плоскости $\eta - \varphi$ меньше, чем параметр R алгоритма anti- k_T .

Ассоциированные треки экстраполируются на поверхность калориметра и классифицируются как треки в конусе струи, если они находятся внутри конуса вокруг оси струи на поверхности калориметра. Треки, которые не попали в конус струи и траектория которых отклонилась за счет магнитного поля детектора CMS, классифицируются как треки вне конуса струи. Импульсы треков далее используются для корректирования энергии ассоциированной струи. Для треков в конусе ожидаемое среднее энерговыделение в калориметре вычитается из энергии струи и добавляется импульс трека, измеренный трековым детектором. Отклик мюонов и электронов в конусе струи учитывается отдельно. В случае с треками вне конуса струи импульс непосредственно прибавляется к струе. Дополнительные коррекции применяются для учета эффективности реконструкции треков в трековом детекторе и для учета эффекта порогов, применяемых к энерговыделению в ячейках адронного калориметра.

2. Данные и Монте-Карло-события

События с рождением Z-бозона и струи (далее для этих событий будет использовано обозначение Z-бозон + струя), а также многоструйные события отбираются из протон-протонных соударений при энергии $\sqrt{s} = 8$ ТэВ (далее данные). Статистика, набранная на установке CMS, соответствует инте-гральной светимости в 19.8 фб⁻¹.

События $Z \to \mu\mu$ отбираются с помощью одиночного мюонного триггера с порогом $p_T > 24$ ГэВ. В последующем анализе требуется наличие двух изолированных мюонов с $p_T > 20$ ГэВ, $|\eta| < 2.4$ и с инвариантной массой двух мюонов в пределах ± 15 ГэВ вокруг номинальной массы Z-бозона. Переменная изоляции мюона вычисляется как сумма поперечных импульсов треков в радиусе $\delta R = 0.4$ вокруг оси мюона, поделенную на поперечный импульс мюона. Значение переменной должно быть меньше 0.12.

Многоструйные события отбираются с помощью двуструйного триггера с требованием, чтобы среднее значение поперечного импульса двух струй с самым большим поперечным импульсом в событии было в определенном интервале. В отобранных событиях поперечный импульс струй находится в интервале от 15 ГэВ до 1 ТэВ.

Монте-Карло-моделирование элементарного взаимодействия Z-бозон + струя было произведено с помощью генератора MadGraph5 [7], моделирующего партонное событие, с последующим генерированием партонных ливней и адронизации с помощью программы Pythia 6.4 [8]. События Z-бозон + + струя также были получены с помощью генераторов элементарного взаимодействия Powheg [9] и Sherpa [10]. Моделирование многоструйных событий было произведено с помощью генераторов Pythia и Herwig [11].

Отклик детектора на Монте-Карло-события был смоделирован с помощью программы Geant4 [12], с последующей имитацией триггеров и реконструкцией. В процессе реконструкции к каждому сигнальному событию было добавлено случайное количество дополнительных событий, причем распределение множественности этих событий совпадает с распределением по множественности дополнительных событий, измеренным по реальным данным.

3. Коррекция энергии струй

Коррекция энергии струй в эксперименте CMS проводится с помощью многоступенчатой процедуры [13–15]:

• на первом этапе коррекции проводится вычитание энергии от дополнительных событий;

 на втором этапе коррекции проводится коррекция абсолютной энергетической шкалы с использованием распределений, полученных с помощью Монте-Карло-событий;

• на третьем этапе коррекции учитывается оставшаяся разница между откликами детектора на струю в данных и Монте-Карло-событиях.

3.1. Первый этап коррекций энергии струй

Первый этап коррекций (далее L1) для Jet-Plus-Track-струй состоит из двух частей: вычитания вклада дополнительных событий и коррекции на эффект порогов на энерговыделение в калориметрических ячейках, которые применяются последовательно к энергии калориметрической струи E_{calo} с помощью мультипликативных факторов c_{PU} и c_{ZSP} соответственно:

$$\begin{split} E_{\text{calo}}^{\text{PU}} &= c_{\text{PU}} E_{\text{calo}}, \\ E_{\text{calo}}^{\text{ZSP,PU}} &= c_{\text{ZSP}} E_{\text{calo}}^{\text{PU}}, \\ E_{\text{JPT}}^{L1} &= E_{\text{JPT}} + \left(E_{\text{calp}}^{\text{ZSP,PU}} - E_{\text{calo}} \right) \end{split}$$

При этом поправка *с*_{PU} вычисляется согласно [13–15] с использованием событий из данных и Монте-Карло с минимальными отборами и без использования отборов. Для оценки поправки используются многоструйные Монте-Карло-события без учета дополнительных событий. Фактор *с*^{MC}_{ZSP} определяется для каждого интервала по псевдобыстроте:

$$c_{\rm ZSP}^{\rm MC}(E_{\rm calo}) = \frac{E_{\rm calo}^{\rm noZSP} - E_{\rm calo}^{\rm ZSP}}{E_{\rm calo}^{\rm ZSP}},$$

где $E_{\rm calo}^{\rm noZSP}$ и $E_{\rm calo}^{\rm ZSP}$ — энергия калориметрической струи, реконструированная с порогом и без порога на энерговыделение в ячейках калориметра.



Рис. 1. Слева: распределение величины коррекции с_{ZSP} для калориметрических струй с поперечной энергией в интервале 50–60 ГэВ и |η| < 0.5. Справа: коррекция как функция поперечной энергии струй с |η| < 0.5

Распределение поправки $c_{\rm ZSP}^{\rm MC}$ показано на рис. 1 для интервала $|\eta| < 0.5$ и поперечного импульса струй в диапазоне 50–60 ГэВ. На этом же рисунке представлена зависимость поправки $c_{\rm ZSP}^{\rm MC}$ от поперечного импульса струи.

Для сравнения поправки на калориметрические пороги между данными и Монте-Карло-событиями была проведена следующая процедура: вклад от дополнительных событий с учетом и без учета калориметрических порогов в конусе калориметрической струи был измерен с использованием данных без калориметрических порогов и без наложения каких-либо триггерных условий. Затем этот вклад был вычтен из энергии струи, измеренной с учетом и без учета калориметрических порогов соответственно, и были построены кривые зависимости поправки на калориметрические пороги от энергии струи, поправленной на вклад от дополнительных событий. Разница в значении для данных и Монте-Карло-событий оказалась меньше 10% для низкоэнергичных струй и менее 1% для струй с $p_T > 80$ ГэВ.

3.2. Относительная и абсолютная шкала энергии струй

Относительная и абсолютная (далее L2L3) шкала струй устанавливается с помощью многоструйных событий и событий с таггирующими частицами (Z + струя, γ + струя). Для этого в Монте-Карло-событиях струи реконструируются из стабильных генераторных частиц (далее под генераторными частицами подразумеваются частицы из Монте-Карло-события, не прошедшие через модель установки, т. е. до преобразования Монте-Карло-события на уровне частиц в отклики детектора CMS посредством модели Geant), исключая мюоны и нейтрино с помощью алгоритма anti- k_T с параметром $R_{anti-k_T} = 0.5$. Данные генераторные струи ассоции-





руются со струями, реконструированными тем же алгоритмом anti- k_T из откликов детектора CMS и полученными посредством модели Geant. Ассоциирование основано на совпадении реконструированных и генераторных струй в плоскости $\eta - \varphi$, причем $R = \sqrt{(\eta_{\text{trk}} - \eta_{\text{jet}})^2 + (\varphi_{\text{trk}} - \varphi_{\text{jet}})^2} < R_{\text{anti-}k_T}$. Отношение поперечного импульса реконструированной и генераторной струи $p_T^{\text{rec}}/p_T^{\text{gen}}$ вычисляется как функция поперечного импульса генераторной струи и псевдобыстроты:

$$p_T^{\text{rec}} = F(p_T^{\text{gen}}, \eta) p_t^{\text{gen}}.$$
 (1)

Уравнение (1) преобразуется, чтобы получить скорректированную энергию струи как функцию

поперечного импульса реконструированной струи и псевдобыстроты:

$$p_T^{\text{gen}} = c^{\text{L2L3}} \left(p_T^{\text{rec}}, \eta \right) p_t^{\text{rec}}$$

Рис. 2 показывает распределение после применения L2L3-коррекции к Монте-Карло-событиям в разных интервалах по псевдобыстроте.

3.3. Различие абсолютной и относительной шкал энергии струй для данных и Монте-Карло

После проведенных коррекций разница между энергетическими шкалами струй в данных и моделировании по-прежнему существует. Для уменьшения этой разницы вводятся дополнительные коррекции, с помощью которых: 1) убирают зависимость по псевдобыстроте в разнице откликов между данными и Монте-Карло-событиями, 2) уменьшают разницу в абсолютных значениях энергии струй в данных и Монте-Карло-событиях.

3.3.1. Дополнительные коррекции по псевдобыстроте

Для получения дополнительных коррекций по псевдобыстроте используется метод асимметрии [15]. Для этого выбираются две струи с наибольшими поперечными импульсами ($p_T^{\rm JPT}$ > > 15 ГэВ).

Струя с наименьшим абсолютным значением псевдобыстроты в пределе $|\eta| < 1.0$ обозначается как эталонная, вторая струя обозначается пробной. Эти две струи должны быть направлены в противоположные стороны в поперечной плоскости детектора $|\varphi_{\text{tag}} - \varphi_{\text{probe}}| > 2.7$. Для классификации событий по дополнительной адронной активности вводится параметр α , который определяется как $\alpha = p_{T,3}/p_{T,1,2 \text{ ave}}$, где $p_{T,3}$ — поперечный импульс

третьей лидирующей струи в событии, а $p_{T,1,2 \text{ ave}}$ средний поперечный импульс двух струй. Чтобы повысить точность вычисления параметра α при определении шкалы Jet-Plus-Track-струй, используются струи, реконструированные с помощью алгоритма Particle Flow [16, 17] (альтернативный алгоритм коррекции энергии струй). Для этого требуется ассоциация эталонной и пробной Jet-Plus-Track-струй с Particle-Flow-струями, а третья Particle-Flow-струя используется для расчета параметра α .

Асимметрия определяется как

$$A = \frac{p_T^{\text{tag}} - p_T^{\text{probe}}}{p_T^{\text{tag}} + p_T^{\text{probe}}}$$

Асимметрия вычисляется для данных и Монте-Карло-событий в интервалах по псевдобыстроте и в инклюзивных интервалах по α : $\alpha < 0.2$, $\alpha < 0.3$, $\alpha < 0.4$, $\alpha < 0.5$. Отношение вычисляется из асимметрии:

$$R = \frac{1-A}{1+A}.$$
 (2)

Распределения по асимметрии строятся в каждом интервале $\eta - \alpha$. Разница в средних значениях асимметрии для данных и Монте-Карло-событий является функцией от разницы откликов детектора на энергию струи в данных и Монте-Карло. Отклик детектора на энергию струи, определенный в уравнении (2), показан на рис. З для данных и Монте-Карло-событий как функция псевдобыстроты для $\alpha < 0.2$. Предполагаемый отклик детектора на энергию струи в точке для данных и Монте-Карло получается путем экстраполяции зависимости отклика струи от α к значению $\alpha = 0$. Дополнительная коррекция вычисляется как отношение откликов



Рис. 3. Средний отклик R для струй в данных и Монте-Карло-событиях как функция псевдобыстроты для α < 0.2 (слева) и дополнительные коррекции по псевдобыстроте как функция псевдобыстроты (справа)

струй для данных и Монте-Карло при для каждого интервала по псевдобыстроте

$$c_{\eta}^{r} = \frac{R^{\text{sim}}(\alpha=0,\eta)}{R^{\text{data}}(\alpha=0,\eta)}$$

Дополнительные коррекции по псевдобыстроте близки к единице в центральной части детектора, которую покрывает трековый детектор, но отклоняются от единицы в передней области. Коррекция применяется только к данным, что позволяет убрать разницу между откликом струй в данных и Монте-Карло событиях для всего диапазона по псевдобыстроте.

3.3.2. Дополнительные коррекции зависимости отклика детектора на струю от поперечного импульса струи

После выравнивания отклика детектора на струю по псевдобыстроте, между данными и Монте-Карло-событиями остается разница в отклике детектора в зависимости от поперечного импульса струи. Дополнительные коррекции по поперечному импульсу получаются путем сравнения событий Z-бозон + струя в данных и в Монте-Карло-событиях. Для получения дополнительных коррекций требуется, чтобы лидирующая струя, реконструированная с помощью алгоритма Jet-Plus-Track с $p_T^{j1} > 15$ ГэВ, была направлена противоположно Z-бозону в поперечной плоскости детектора $|\varphi^Z - \varphi^{j1}| > 2.7$ и чтобы она совпадала в плоскости η - φ с Particle-Flow-струей, обладающей *p*_T > 10 ГэВ. Параметр *α*, который характеризует наличие дополнительных струй в событии, определяется с использованием струй, реконструированных с помощью алгоритма Particle-Flow. Параметр α определяется как $\alpha = p_T^{j2}/p_T^Z$, где p_T^{j2} — поперечный импульс второй по величине PF-струей в событии. Распределения отклика струи $R = p_T^{\prime 1} / p_T^Z$, проинтегрированные по области $p_T^Z > 50$ ГэВ, рассматриваются в интервалах по α $(0.1 < |\alpha| < 0.15, 0.15 < |\alpha| < 0.2$ ит.д.) для данных и Монте-Карло-событий. Так как на предыдущей стадии отклик струи был выравнен по псевдобыстроте, то для абсолютной шкалы достаточно использовать события, в которых лидирующая струя находится в центральной ($|\eta| < 1.3$) области детектора. Отклик R аппроксимируется с помощью распределения Гаусса, и среднее значение отклика струи строится в зависимости от а. Отношение зависимости отклика от для данных и Монте-Карло-событий хорошо параметризуется полиномом первого порядка, а дополнительные коррекции по поперечному импульсу определяются как

$$c^{R} = \frac{R^{\rm sim}(\alpha = 0)}{R^{\rm data}(\alpha = 0)}$$

Данные коррекции применяются к данным во всем диапазоне по псевдобыстроте после применения дополнительных коррекций по псевдобыстроте.

3.4. Неопределенность энергетической шкалы струи

Для вычисления систематической неопределенности энергетической шкалы для Jet-Plus-Track-струй были использованы события со струями. При этом требовалось, чтобы две самые энергичные Jet-Plus-Track-струи совпадали по направлению с двумя самыми энергичными Particle-Flow-струями. Было установлено ограничение на параметр $\alpha = p_{T3}/p_{T1,2}$ ave < 0.2, а также построены распределения:

- еления: $p_T^{\text{PF probe}}/p_T^{\text{PF tag}}$ в зависимости от $p_T^{\text{PF tag}}$; $p_T^{\text{JPT probe}}/p_T^{\text{PF tag}}$ в зависимости от $p_T^{\text{PF tag}}$; $p_T^{\text{PF probe}}$ в зависимости от $p_T^{\text{PF tag}}$.

Используя формулы, приведенные выше, значение отношения $p_T^{\text{JPT probe}}/p_T^{\text{PF probe}}$ получили в интервалах по псевдобыстроте как функцию $p_{\tau}^{\mathrm{PF\,probe}}$. Полученные зависимости были параметризованы логарифмической функцией, которая использовалась для переноса систематических ошибок, полученных для Particle-Flow-струй к Jet-Plus-Track-струям. Разница в неопределенности между Particle-Flowи Jet-Plus-Track-струями составляет менее 1%.

3.5. Энергетическое разрешение струй

Энергетическое разрешение струй вычисляется путем анализа многоструйных событий при использовании метода баланса поперечного импульса двух струй. Оно соотносится с разрешением асимметрии А (см. разд. 4.3.1) как

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{2} \,\sigma_A.$$

Для отбора событий используются Particle-Flowструи с требованием соответствия двух лидирующих Particle-Flow-струй с лидирующими Jet-Plus-Track-струями. Параметр α вычисляется по третьей по поперечному импульсу Particle-Flow-струе как $\alpha = p_{T3}^{\rm PF}/p_{T12 \, \rm ave}^{\rm PF}$. Стандартное отклонение асимметрии вычисляется в интервалах по среднему поперечному импульсу двух лидирующих Jet-Plus-Track-струй как функция параметра. При этом обе струи должны находиться в одном и том же интервале по абсолютному значению псевдобыстроты. В каждом интервале по поперечному импульсу и по псевдобыстроте находится значение при параметре $\alpha = p_{T3}^{\rm PF}/p_{T12 \, \rm ave}^{\rm PF} \rightarrow 0.$ (см. разд. 4.3.1.). Процедура применяется для данных и Монте-Карло-событий. Для учета эффектов, связанных с естественным дисбалансом двуструйных событий (глюонная радиация), эта же процедура повторяется для генераторных струй.

Зависимость разрешения от поперечного импульса показана на рис. 4. Разрешение для данных на 10-20% хуже, чем для Монте-Карло-событий. Также на рисунке показано «истинное» разрешение (стандартное отклонение распределения $p_T^{\text{rec}}/p_T^{\text{gen}}$), полученное из Монте-Карло-событий. Оно согласуется с разрешением, полученным методом асимметрии.



Рис. 4. Энергетическое разрешение струи в зависимости от p_T для $|\eta| < 0.5$ (слева) и $0.5 < |\eta| < 1.1$ (справа) для струй в данных и Монте-Карло-событиях. Также приводится истинное разрешение, полученное для Монте-Карло-событий

Идентификация струй из сигнального события Определение струи из сигнального и дополнительного событий

Процедура разделения струй из сигнальных и дополнительных событий была разработана и протестирована с помощью Монте-Карло-событий. Разделение реконструированных струй в Монте-Карло-событиях на те, что относятся к сигнальному событию, и на те, что происходят из дополнительных



Рис. 5. Распределение по поперечному импульсу струй из сигнального и дополнительного события для Монте-Карло-событий Z-бозон + струя при $p_T > 20$ ГэВ и $|\eta| < 2.6$

событий, основано на информации о смоделированном сигнальном взаимодействии. Для этого генераторные струи формируются из стабильных частиц, принадлежащих сигнальному событию, исключая мюоны и нейтрино. Процедура происходит с помощью алгоритма anti- k_T с радиусом R = 0.5. Далее вычисляется расстояние между осями генераторной и реконструированной струй в плоскости $\eta - \varphi$. Если расстояние *R* между реконструированной струей и генераторной струей в плоскости $\eta - \varphi$ меньше 0.2, то такая струя называется струей из сигнального события. В случае R > 0.4 струя называется струей из дополнительного события. Распределение по поперечному импульсу струй из сигнального и дополнительного события для Монте-Карло событий Z-бозон + струя показано на рис. 5.

4.2. Идентификации струй из дополнительных событий

Для разделения струй из сигнальных и дополнительных событий в области используются следующие переменные:

 полное число событий (сигнальное + дополнительные);

 значение поперечного импульса и псевдобыстроты струи;

 множественность калориметрических башен, являющихся частью струй;

• множественность треков, идущих от сигнального события и ассоциированных со струей (см. описание алгоритма Jet-Plus-Track в разд. 2);

переменная, определяемая как

$$\beta = \frac{\sum\limits_{i \in jet} p_{T,i}}{\sum\limits_{i \in track} p_{T,i}},$$
(3)

где $\sum p_{T,i}$ — сумма поперечных импульсов треков,

которые идут от сигнального события и проассоциированы со струей в диапазоне $|\eta| < 2.6$. Сумма ∑ *p*_{*T,i*} в уравнении (3) − это сумма попереч*i*∈track

ных импульсов треков, идущих из любого события, но попадающих в конус струи на поверхности калориметра (см. рис. 6);



вершина (сигнальная) вершина

Рис. 6. Схематическое изображение треков в струе

ширина струй вдоль главных осей η и φ :

$$A_{\eta}^{\text{trk(calo)}} = \frac{\sum_{i \in jet} \left(\eta_{i}^{\text{trk(calo)}} - \eta^{jet}\right)^{2} p_{T,i}^{\text{trk(calo)}}}{\sum_{i \in jet} p_{T,i}^{\text{trk(calo)}}}, \qquad (4)$$
$$\sum_{i \in jet} \left(\varphi_{i}^{\text{trk(calo)}} - \varphi^{jet}\right)^{2} p_{T,i}^{\text{trk(calo)}}$$

$$A_{\varphi}^{\text{trk(calo)}} = \frac{i \in \text{jet}^{(T_{i}, T_{i})} + 1, i}{\sum_{i \in \text{jet}} p_{T,i}^{\text{trk(calo)}}}.$$
 (5)

В диапазоне $2.6 < |\eta| < 4.7$ детектор не измеряет треки заряженных частиц и поэтому используется другой набор переменных, использующих только калориметрическую информацию:

• полное число событий (сигнальное + дополнительные);

• значение поперечного импульса и псевдобыстроты струи;

• множественность калориметрических башен. являющихся частью струй;

и A_{η}^{calo} и $A_{\varphi}^{\text{calo}}$ (см. (4), (5)). Переменные, описанные выше, используются в качестве входных данных для алгоритма машинного обучения Boosted Decision Tree (BDT) [18]. Переменные, поданные на вход классификатора BDT, преобразуются в одну переменную (PUID), которая имеет разные распределения для сигнальных

струй и струй из дополнительных событий. Обучение классификатора BDT происходит отдельно для двух диапазонов по псевдобыстроте: внутри области покрытия трекового детектора ($|\eta| < 2.6$) и вне него $(2.6 < |\eta| < 4.7)$. Обучение проводится на многоструйных Монте-Карло-событиях и событиях Z-бозон + струя, причем заведомо известно, происходит ли струя из сигнального или дополнительного событий. После обучения классификатор выдает вероятность того, что конкретная струя принадлежит сигнальному событию. Значение переменной PUID на выходе классификатора тем выше,



Puc. 7. Распределение PUID для струй из сигнальных и дополнительных событий с $p_T > 20$ ГэВ и $|\eta| < 2.6$ для Монте-Карло-событий Z-бозон + струя



Рис. 8. Распределение PUID для струй с *p_T* > 20 ГэВ и $|\eta| < 2.6$ из данных и Монте-Карло-событий Z -бозон + струя

чем выше вероятность того, что струя принадлежит сигнальному событию. Распределение PUID для струй из сигнальных и дополнительных событий с $p_T > 20$ ГэВ и $|\eta| < 2.6$ для Монте-Карло-событий Z-бозон + струя показано на рис. 7.

Распределения PUID для струй с $p_T > 20$ ГэВ и $|\eta| < 2.6$ из данных и Монте-Карло-событий Z-бозон + струя показаны на рис. 8. Видно хорошее совпадение результатов между данными и Монте-Карло-событиями.

4.3. Ограничения на параметр PUID для алгоритма выделения сигнальных струй

Для разделения струй от сигнальных и дополнительных событий используется критерий PUID > PUID^{cut}. Значение критерия PUID^{cut} лежит в пределах $-1 < PUID^{cut} < 1$. На рис. 9 показана эффективность отбора струй от сигнальных и дополнительных событий как функция PUID^{cut} для струй с $p_T > 20$ ГэВ, $|\eta| < 2.6$ и для струй с $p_T > 25$ ГэВ, $2.6 < |\eta| < 4.7$ для Монте-Карло-событий *Z*-бозон + струя.



Рис. 9. Эффективность отбора истинных струй и струй из дополнительных событий как функция PUID^{cut} для струй с $p_T > 20$ М и для струй с $p_T > 25$ ГэВ для Монте-Карло-событий Z-бозон + струя

Определены три рабочие точки, «слабая», «средняя» и «сильная», для которых эффективность отбора сигнальных струй составляет 85, 90 и 95% соответственно в области $|\eta| < 2.6$ для струй с $p_T > 20$ ГэВ для Монте-Карло-событий Z-бозон + струя. Для области $2.6 < |\eta| < 4.7$ выбраны рабочие точки, которые соответствуют эффективности отбора 70, 80 и 90%. В табл. 1, 2 показаны значения PUID^{cut}.

На рис. 9 показана эффективность отбора для струй из дополнительных событий (слева) и сиг-

Таблица 1

Рабочие точки для струй с $P_T > 20$ ГэВ и $|\eta| < 2.6$ для Монте-Карло-событий Z-бозон + струя

Рабочая точка	PUID ^{cut}	Эффективность идентификации струй из сигнального события, %	Эффективность идентификации струй из дополнительного события, %
Слабая	0.3	95	9
Средняя	0.7	90 4	
Сильная	0.9	80	1.5

Таблица 2

Рабочие точки для струй с $P_T > 20$ ГэВ и $2.6 < |\eta| < 4.7$ для Монте-Карло-событий Z-бозон + струя

Рабочая точка	PUID ^{cut}	Эффективность идентификации струй из сигнального события, %	Эффективность идентификации струй из дополнительного события, %
Слабая	-0.07	90	75
Средняя	0.12	80	55
Сильная	0.25	70	42

нальных струй (справа) как функция абсолютного значения псевдобыстроты струй для трех рабочих точек, описанных выше. Для струй используются Монте-Карло-события Z-бозон + струя. Сравнения Монте-Карло-событий с данными осуществляется на Монте-Карло-событиях с процессом Дрелла-Яна отдельно для лидирующей и сублидирующей струи. Наблюдаемая разница лежит в диапазоне 5–10%.

5. Разделение кварковых и глюонных струй

5.1. Определение типа струи

Аромат струи определяется с помощью анализа информации, полученной из Монте-Карло-событий, в частности информации о партоне, инициирующем струю. Партон может быть глюоном или кварком. Для идентификации аромата струи на уровне Монте-Карло-событий берутся струи из сигнального события со «слабым» критерием отбора, описанным в предыдущем разделе. Далее выбирается партон с $p_T > 15$ ГэВ, который ассоциируется со струей, в случае если их оси в плоскости η - φ находятся на меньшем расстоянии, чем параметр реконструкции струи $R_{\text{anti-}k_T}$.

5.2. Параметры кварк-глюонного дискриминатора

Алгоритм разделения кварковых и глюонных струй разработан для струй с $|\eta| < 2.6$. Он использует переменные, которые применяются для выделения сигнальных струй (см. разд. 4): множественность треков, ассоциированных со струей: A_{η}^{calo} , A_{φ}^{calo} ,

 A_{η}^{trk} , A_{φ}^{trk} . Также вводится переменная p_T^D , которая вычисляется из поперечных импульсов треков:

$$p_T^D = \sum_{i \in \text{jet}} p_{T,i}^2 / \sum_{i \in \text{jet}} p_{T,i}.$$

Переменные, описанные выше, используются в качестве входных данных для машинного обучения. Обучение классификатора BDT происходит отдельно для четырех интервалов по поперечному импульсу струи, 30–50, 50–100, 100–300, 300–600 ГэВ. Для диапазона 30–50 ГэВ были смоделированы Монте-Карло-события Z-бозон + струя с помощью генератора MADGRAPH5. Для остальных диапазонов использовались Монте-Карло-события со струями, полученные с помощью генератора РҮТНІА6. После обучения классификатор выдает вероятность того, что конкретная струя является кварковой или глюонной. Значение выходной переменной (QGID) на выходе классификатора тем выше, чем выше вероятность струи являться кварковой. Для моделирования Монте-Карло-событий Z-бозон + струя также использовались генераторы POWHEG и SHERPA. Было установлено, что результаты всех генераторов согласуются с точностью 10–15% по средним значениям переменных, описанных выше.



Рис. 10. Распределение QGID для разного количества полного числа событий. Слева показано распределение для струй в диапазоне 30 < *p*_T < 50 ГэВ, где применяются коррекции на переменные дискриминатора, связанные с дополнительными событиями. Справа показано распределение QGID для струй 50 < *p*_T < 100 ГэВ без коррекции на дополнительные взаимодействия



Рис. 11. Сравнение распределений переменной QGID для данных и Монте-Карло-событий *Z*-бозон + струя. На рисунке слева моделирование проведено с помощью генератора MADGRAPH5, на рисунке справа — генератором POWHEG

Переменные дискриминатора для струй с $p_T < 50$ ГэВ чувствительны к числу дополнительных событий. В этом случае для переменных, используемых для разделения струй, определяются коррекции как функция общего числа событий. На рис. 10 показано распределение QGID для разного общего числа событий. Слева на рис. 10 показано распределение для струй с $30 < p_T < 50$ ГэВ с применением коррекции на число дополнительных событий. Справа на рис. 10 показано распределение дополнительных событий. Справа на рис. 10 показано распределение QGID для струй $50 < p_T < 100$ ГэВ без коррекции на число дополнительных событий.

На рис. 11 представлено сравнение распределений переменной QGID для данных и Монте-Карло-событий Z-бозон + струя. На рисунке слева моделирование Монте-Карло-событий Z-бозон + струя было произведено с помощью генератора MADGRAPH5, на правом рисунке — с помощью генератора POWHEG. Результаты двух генераторов согласуются с высокой точностью.

> Таблица 3 Эффективность и чистота разделения кварков и глюонов

		Jet-Plus-Track	Particle-Flow
Кварковые	эффективн.	0.61%	0.62%
струи (uds)	чистота	0.75%	0.76%
Глюонные	эффективн.	0.50%	0.51%
струи	чистота	0.49%	0.51%

В табл. З показана эффективность и чистота разделения кварковых и глюонных Jet-Plus-Track-струй с $30 < p_T < 50$ ГэВ для Монте-Карло-событий Z-бозон + струя. Отбор осуществлялся со следующими условиями: QGID > 0.4 для кварковых струй и QGID < -0.3 для глюонных струй. Результаты кварк-глюонного дискриминатора для Particle-Flow-струй практически не отличаются от результатов с Jet-Plus-Track-струями. Для Particle-Flow-струй были установлены следующие пороги: QGID > 0.6 для кварковых струй и QGID < 0.2 для глюонных струй.

Заключение

На основе данных детектора CMS при энергии $\sqrt{s} = 8$ ТэВ на статистике, соответствующей интегральной светимости в 19.8 фб⁻¹, была установлена энергетическая шкала для струй, реконструированных алгоритмом Jet-Plus-Track. Также была проведена оценка разрешения по энергии струи в зависимости от поперечного импульса струи. Было получено, что разрешение по энергии Jet-Plus-Track-струй сравнимо с разрешением по энергии струй, реконструированных алгоритмом Particle Flow. Алгоритм был дополнен процедурами разделения струй из сигнальных и дополнительных событий, а также процедурой разделения кварковых и глюонных струй. Полученные результаты позволяют использовать алгоритм для анализа данных детектора CMS в качестве дополнительного алгоритма. При близости величины разрешения и систематических ошибок при реконструкции Jet-Plus-Track- и Perticle-Flow-струй используются разные подходы при комбинировании информации с разных субдетекторов CMS, что особенно важно при проведении физического анализа.

Список литературы

- 1. *CMS Collaboration*. Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS. CMS Physics Analysis Summary CMS-PAS-JME-09-002. 2009.
- 2. *Green D.* et al. Energy Flow Objects and Usage of Tracks for Energy Measurement in CMS. CMS Note 2002/036, 2002.
- 3. *CMS Collaboration* // Eur. Phys. J. C. 2005. **40**, N 5. P. 33.
- 4. CMS Collaboration // JINST. 2008. 3. S08004.
- 5. *CMS Collaboration //* Pileup Jet Identification. CMS Physics Analysis Summary CMS-PAS-JME-13-005. CERN, Geneva, 2013.
- Cacciari M., Salam G.P., Soyez G. // JHEP. 2008. 04. P. 063. arXiv: 0802.1189.
- Stelzer T., Long W. // Comput. Phys. Commun. 1994.
 81. P. 57. arXiv: hep-ph/9401258.
- Sjostrand T., Mrenna S., Skands P.Z. // JHEP. 2006.
 05. P. 026. arXiv: hep-ph/0603175.
- 9. Frixione S., Nason P., Ridolfi G. The POWHEG-hvq manual. Ver. 1.0. arXiv: 0707.3081.
- 10. Gleisberg T. et al. // JHEP. 2009. 02. P. 007.
- 11. Bahr M. et al. // Eur. Phys. J. 2008. C58. P. 639.
- Agostinelli S. et al. // Nucl. Instr. Meth. 2003. A506.
 P. 250.
- 13. *CMS Collaboration*. Jet Energy Corrections Determination at $\sqrt{s} = 7$ TeV. CMS Physics Analysis Summary CMS-PAS-JME-10-010. CERN, Geneva, 2010.
- 14. *CMS Collaboration*. Jet Resolution Determination at $\sqrt{s} = 7$ TeV. CMS Physics Analysis Summary CMS-PAS-JME-10-014. CERN, Geneva, 2010.
- 15. CMS Collaboration // JINST. 2008. 3. S08004.
- CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. CMS Physics Analysis Summary CMS-PAS-PFT-09-001. CERN, Geneva, 2009.
- 17. *CMS Collaboration*. Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from *pp* Collisions at 7 TeV. CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.
- Hocker A. TMVA Toolkit for Multivariate Data Analysis. PoS ACAT. 2007. 040. arXiv: physics/0703039.

The operating principles of an algorithm for jet reconstruction using a calorimeter and trackers in collisions of high-energy particles

O.L. Kodolova^{1,a}, S.V. Obraztsov^{1,b}, I.N. Vardanyan^{1,c}, A.N. Nikitenko², N.V. Lichkovskaya², V.B. Gavrilov², M.D. Toms², CMS Collaboration

¹Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University. Moscow 119991, Russia. ²Institute of Theoretical and Experimental Physics, National Research Center «Kurchatov Institute». Moscow 117218, Russia.

E-mail: ^a olga.kodolova@cern.ch, ^b stepan.obraztsov@cern.ch, ^c cirina.vardanyan@cern.ch.

On the basis of data on proton-proton collisions at the energy $\sqrt{s} = 8$ TeV and with the integrated luminosity of 19.8 fb⁻¹, the resolution of the transverse momentum and the uncertainty of the energy scale were estimated for jets that were reconstructed using the Jet-Plus-Track algorithm. The performance of the algorithm for the identification of jets that originated from the signal event and the separation of jets into quark and gluon jets was also assessed.

Keywords: jet reconstruction, jet calibration, high-energy physics.
PACS: 13.85.-t. *Received 10 May 2017*.
English version: *Moscow University Physics Bulletin*. 2018. **72**, No. 3. Pp. 241–251.

Сведения об авторах

- 1. Кодолова Ольга Леонидовна канд. физ.-мат. наук, вед. науч. сотрудник; e-mail: olga.kodolova@cern.ch.
- 2. Образцов Степан Владимирович мл. науч. сотрудник; e-mail: stepan.obraztsov@cern.ch.
- 3. Варданян Ирина Николаевна канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: irina.vardanyan@cern.ch.
- 4. Никитенко Александр Николаевич канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: alexander.nikitenko@cern.ch.
- 5. Личковская Наталья Владимировна канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: natalia.lychkovskaya@cern.ch.
- 6. Гаврилов Владимир Борисович доктор физ.-мат наук, зав. лабораторией; e-mail: Vladimir.gavrilov@cern.ch.
- 7. Томс Мария Дмитриевна инженер, e-mail: maria.tarkhova@cern.ch.