

## Разделение кварковых и глюонных струй в соударениях тяжелых ионов при высоких энергиях

С. В. Образцов,<sup>1,\*</sup> О. Л. Кодолова<sup>1,†</sup>

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

(Поступила в редакцию 28.05.2023; после доработки 09.06.2023; принята к публикации 12.07.2023)

Исследована возможность разделения кварковых и глюонных струй в соударениях тяжелых ионов как при наличии эффекта гашения струй, так и при его отсутствии. С помощью методов машинного обучения разработана методика восстановления фракции кварковых и глюонных струй, с использованием переменных, характеризующих форму, множественность и заряд струи. Показана возможность использования данных переменных для исследования характеристик плотной среды, возникающей при столкновении ионов высокой энергии

PACS: 13.85.-t УДК: 539.1.07

Ключевые слова: физика тяжелых ионов, адронные струи, методы машинного обучения.

DOI: [10.55959/MSU0579-9392.78.2350203](https://doi.org/10.55959/MSU0579-9392.78.2350203)

### ВВЕДЕНИЕ

Энергичные партоны, образующиеся в результате жестких процессов при соударениях тяжелых ионов высоких энергий, теряют энергию при прохождении плотной среды, кварк-глюонной плазмы [1, 2]. Механизмы этих потерь энергии до конца не известны. В результате партонного ливня и последующей адронизации партоны наблюдаются в эксперименте как поток энергичных частиц, вылетающих в узком конусе (струя). Причем ширина конуса зависит от поперечного импульса партона, а также его типа  $u$ ,  $d$ ,  $s$ ,  $c$ ,  $b$  (кварк) или  $g$  (глюон). Важной характеристикой струи является ее заряд, так как он отражает заряд партона, инициирующего струю. Соотношение числа кварков и глюонов в выборке событий зависит от типа жесткого процесса. Например, в событиях  $Z$ +струя соотношение кварковых и глюонных струй 70% кварковых к 30% глюонных струй, в то время как для многоструйных событий соотношение обратное [3]. Потери энергии партонов проявляются в эксперименте в виде подавления выхода энергичных адронов и струй с большим поперечным импульсом, а также модификации формы струй. Совокупность данных эффектов принято называть эффектом гашения адронных струй [4–7], который является одним из основных сигналов возникновения кварк-глюонной плазмы. Поскольку развитие партонного ливня у кварков и глюонов несколько отличается, то в результате получаются струи разной ширины. Струи, инициированные глюонами, шире, чем струи, инициированные легкими кварками [8]. При прохождении плотной среды не исключено дополнительное

уширение струй, причем степень уширения может зависеть от типа партона и от характеристик плотной среды. В ряде теоретических моделей ожидается, что и доля глюонных (кварковых) струй может существенно меняться при возникновении кварк-глюонной плазмы и измерение фракции глюонных (кварковых) струй в зависимости от центральности событий может являться важной экспериментальной наблюдаемой для проверки теоретических моделей [9]. Поэтому возможность разделения кварковых и глюонных струй, а также изучение изменения характеристик струй при прохождении партонов через плотную среду, является важной задачей для регистрации возникновения кварк-глюонной плазмы, а также изучения ее характеристик.

Для разделения струй, инициируемых легкими кварками ( $u$ ,  $d$ ,  $s$ ) от глюонных струй в  $pp$ -взаимодействиях используется ряд наблюдаемых, характеризующих форму струи и множественность частиц в струе и разработаны соответствующие методики на основе машинного обучения [10]. В отличие от протон-протонных соударений в соударениях тяжелых ионов существуют дополнительные сложности, связанные с фоном от мягкой компоненты, зависимостью переменных от прицельного параметра соударения (от центральности события) и от эффектов прохождения струй через кварк-глюонную плазму. Поэтому процедура разделения кварковых и глюонных струй должна быть дополнена выделением наблюдаемых, характеризующих струю, из фона от тяжелоионного взаимодействия.

Поскольку энергия струй уменьшается в результате гашения и характеристики струй меняются при прохождении через плотную среду и в тоже время зависят от поперечного импульса струи, то для сравнения характеристик струй в плотной среде с  $pp$ -взаимодействиями или изучения зависимости от прицельного параметра необходимо вы-

\* E-mail: [stepan.obraztsov@cern.ch](mailto:stepan.obraztsov@cern.ch)

† E-mail: [olga.kodolova@cern.ch](mailto:olga.kodolova@cern.ch)

брать процесс с той же самой жесткостью взаимодействия. Иными словами, процесс должен характеризоваться частицей, которая не чувствительна к плотной среде, например, процесс фотон плюс струя или Z бозон плюс струя. Тогда зависимость характеристик струи от прицельного параметра события можно изучать в зависимости от поперечной энергии фотона.

## 1. МОДЕЛИРОВАНИЕ СОУДАРЕНИЙ ЯДЕР СВИНЦА

Для моделирования жесткого процесса фотон+струя в столкновениях тяжелых ионов использовался Монте–Карло генератор PYQUEN, в то время как событие с тяжелыми ионами (фон от мягкой компоненты события) моделировалось отдельно с помощью Монте–Карло генератора HYDJET [11]. События моделировались с разными прицельными параметрами с учетом и без учета эффекта гашения струй. При смешивании сигнального события и мягкой компоненты выбирались события с одним и тем же прицельным параметром. Для последующей обработки на частицы события накладывались отборы, соответствующие геометрии детектора CMS [12].

Диаграммы Фейнмана для рождения прямых фотонов показаны на рис. 1. Фотоны рождаются в результате комптоновского рассеяния кварка на глюоне, а также при аннигиляции пары кварк-антикварк. Монте–Карло генератор PYQUEN считает в качестве входных данных прицельный параметр фонового события HYDJET. Сигнальные события производятся как с эффектом гашения струй, так и без него. Выбор сигнального процесса обусловлен тем, что фотон используется как тагирующая частица, которая после рождения проходит через кварк-глюонную плазму без потерь и в то же время характеризует жесткость взаимодействия и, соответственно, поперечный импульс партона до прохождения плотной среды. После моделирования сигнала два события смешиваются, в результате чего получается событие, где, как и в реальных данных, есть мягкая компонента и жесткая компонента.

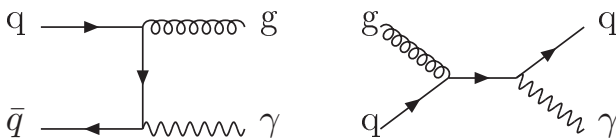


Рис. 1. Диаграммы Фейнмана для рождения прямых фотонов. Комптоновское рассеяния кварка на глюоне (слева), аннигиляция пары кварк-антикварк (справа)

Для анализа использовались два набора данных с и без эффекта гашения струй. Для получения равномерной статистики в зависимости от  $P_T$  фотона жесткое событие моделируется в интервалах 80–120, 120–170 и 170–500 ГэВ по поперечному им-

пульсу фотона по 500 тысяч событий в каждом диапазоне, что дает итоговую статистику в 1.5 миллиона событий без эффекта гашения струй и столько же с эффектом гашения струй.

## 2. РЕКОНСТРУКЦИЯ АДРОННЫХ СТРУЙ

Кластеризация адронных струй алгоритмом anti- $k_T$  [13] с параметром радиуса 0.4 производится с помощью пакета программного обеспечения FastJet [14]. Перед тем как произвести кластеризацию, необходимо осуществить вычитание фона от тяжелоионного события-подложки. Для этого был применен итерационный метод, описанный в [15]. Пространство  $\eta, \phi$  делится на кольца толщиной  $\Delta\eta = 0.1$ , что приблизительно соответствует сегментации калориметров CMS. Затем из энергии каждой частицы вычитается среднее значение энергии по кольцу. После этого производится кластеризация адронных струй. На второй итерации среднее значение по кольцу  $\eta$  вычисляется только с учетом частиц, которые не кластеризовались в струю.

## 3. ХАРАКТЕРИСТИКИ КВАРКОВЫХ И ГЛЮОННЫХ СТРУЙ В ТЯЖЕЛОИОННОМ СОБЫТИИ

Для разделения кварковых и глюонных струй с использованием методов машинного обучения необходимо подобрать переменные, которые принимают разные значения для кварковых и глюонных струй. Дополнительным требованием является возможность вычитания фоновой компоненты. Изначально рассматривались классические переменные, используемые в pp-соударениях [16]. Дополнительно были включены в рассмотрение ширина и заряд струи.

Для вычитания фоновой компоненты из наблюдаемых характеристик струй используется метод случайного конуса [17]. Метод состоит в вычислении значения переменных для конуса с радиусом  $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$  и осью  $\eta, \phi$ , зеркально отраженной относительно оси рассматриваемой струи. При этом значения переменных вычисляются только для частиц, принадлежащих тяжелоионному событию, по аналогии с тем, как в реальном эксперименте он вычисляется на данных, взятых с минимальными отборами. Фоновую компоненту события можно вычитать пособытийно для каждой струи либо можно вычитать для каждой переменной среднее значение фоновой компоненты для данного значения центральности. Оба метода работают, но первый подвержен флуктуациям для низкоэнергичных струй. Второй метод можно реализовать путем аппроксимации полиномом четвертой степени зависимости среднего значения переменной от центральности соударения. Для оценки качества вычитания фоновой компоненты, струи после вычитания фона сравниваются с «истинными» струями, которые ре-

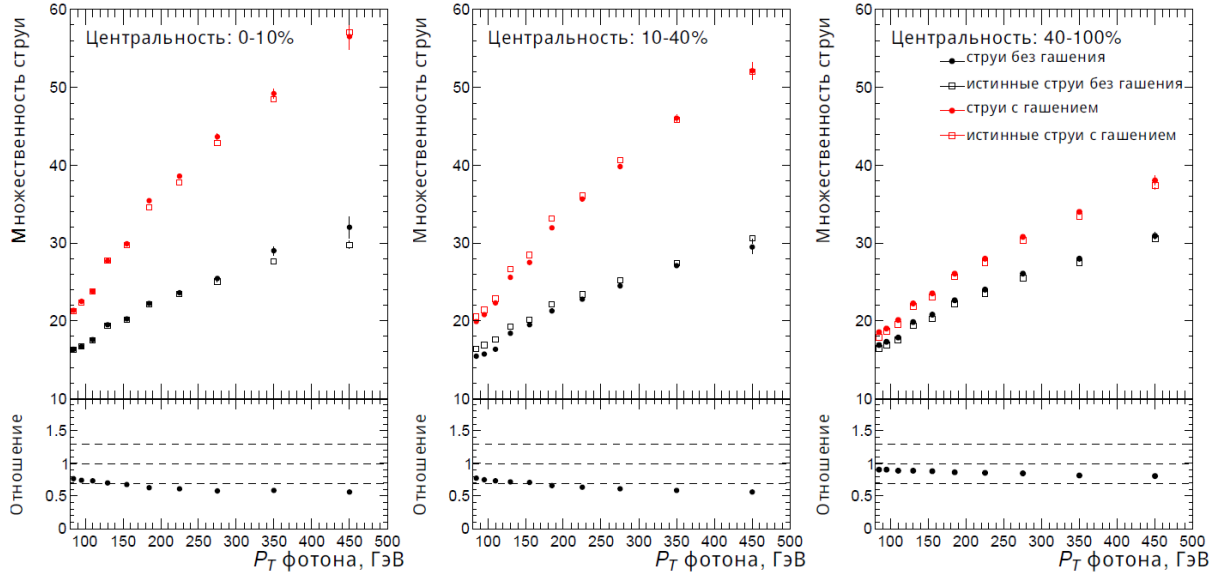


Рис. 2. Измеренная множественность струи после вычитания фона от события-подложки в зависимости от  $P_T$  фотона в сравнении с истинным значением для трех интервалов по центральности. Приведены зависимости для вариантов с гашением и без гашения сигнальной струи

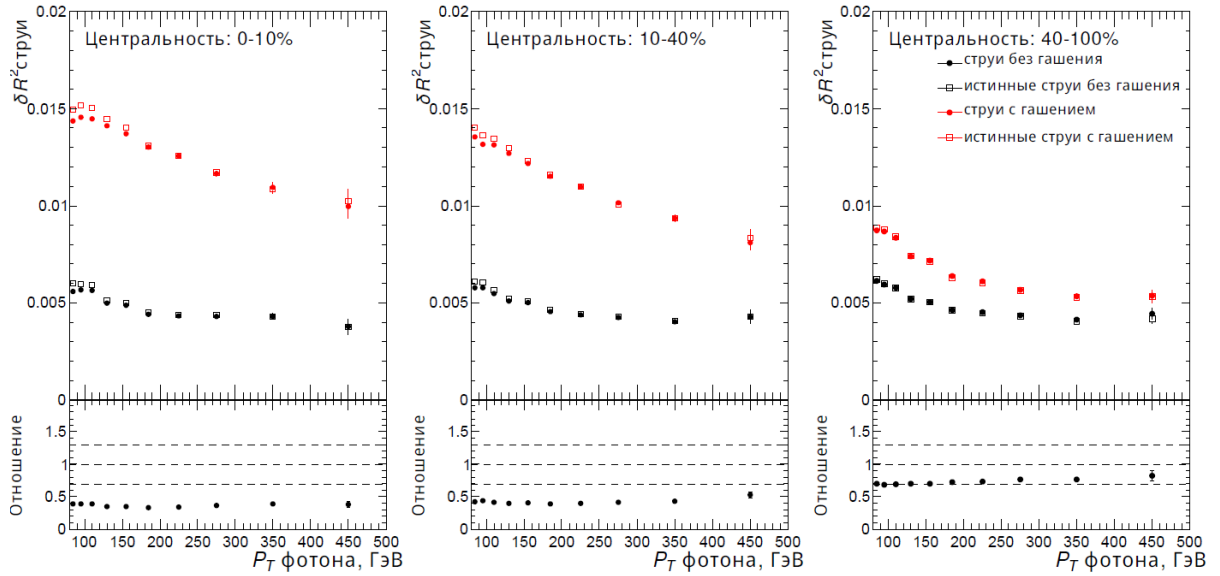


Рис. 3. Измеренная ширина струи после вычитания фона от события-подложки в зависимости от  $P_T$  фотона в сравнении с истинным значением для трех интервалов по центральности. Приведены зависимости для вариантов с гашением и без гашения сигнальной струи

конструируются только с использованием сигнальных событий.

Множественность заряженных частиц в струе вычисляется как количество заряженных частиц, попадающих в конус вокруг оси струи с радиусом 0.4. Для данной переменной характерен рост среднего значения с ростом  $P_T$  фотона. Также можно наблюдать отсутствие зависимости от центральности для событий без эффекта гашения струй. Разница между переменными с или без эффекта гашения струй увеличивается в случае более центральных событий, как показано на рис. 2.

Средняя ширина струи вычисляется по формуле

$$\langle \delta R^2 \rangle (P_T) = \frac{\sum_{i \in jet} \Delta R_i^2 P_T^i}{\sum_{i \in jet} P_T^i}, \quad (1)$$

где  $\Delta R^2 = \Delta \eta^2 + \Delta \phi^2$  — это квадрат расстояния от частицы до оси струи в пространстве  $\eta, \phi$ . С ростом  $P_T$  струя сужается. В случае наличия эффекта гашения струй также наблюдается уширение струи по мере увеличения центральности соударений. Распределения показаны на рис. 3.

Заряд струи вычисляется по формуле

$$Q(P_T) = \frac{1}{\sqrt{P_T(jet)}} \sum_{i \in jet} q(i) \sqrt{P_T(i)} \quad (2)$$

Заряд струи остается постоянным вне зависимости от центральности или  $P_T$  фотона. Заряд струи уменьшается при наличии эффекта гашения струй. Распределения показаны на рис. 4.

Множественность заряженных частиц в струе, ширина заряд струи были исследованы отдельно для кварковых и глюонных струй на всем диапазоне центральности соударений, как в случае с гашением струй, так и без него (рис. 5).

#### 4. МЕТОДИКА РАЗДЕЛЕНИЯ КВАРКОВЫХ И ГЛЮОННЫХ СТРУЙ

Для разработки кварк-глюонного дискриминатора был использован метод машинного обучения Gradient Boosted Decision Tree (GBDT) из пакета ROOT TMVA [18]. Данный метод применяется похожим образом в задаче по выделению струи из сигнальной вершины в pp-соударениях [19]. Выбор в пользу деревьев решений был сделан по причине простоты их применения, а также низких требований к вычислительным мощностям. Альтернативой деревьям решений считаются глубокие нейронные сети (Deep Neural Networks), которые являются популярным методом в идентификации b и c струй. Глубокие нейронные сети используют входные переменные низкого уровня или же вовсе работают с двумерными «изображениями» струй. Данная методика интенсивно внедряется в работе с pp-соударениями, однако находится за пределами рассмотрения в данной работе.

Для обучения классификатора GBDT используются три переменные, описанные выше. При обучении заранее известно, какая струя является глюонной, а какая — кварковой, 90% данных используются для обучения, 10% для оценки результатов. Обучение проводится отдельно для трех разных диапазонов по центральности 0–10%, 10–30%, 30–100% и отдельно для данных с эффектом гашения струй и без него. Применяться обученные модели будут в тех же диапазонах. Критерием оценки эффективности обучения является определение долей кварковых и глюонных струй и их сравнение с истинными значениями из Монте-Карло. Для выбора рабочей точки следует построить распределения зависимости эффективности определения типа струи от отбора по выходному параметру GBDT, при этом разность эффективности выделения кварковых и глюонных струй должна быть максимальна. Эффективность для глюонных и кварковых струй, а также разница эффективностей показаны на рис. 6. Видно, что оптимальный порог по значению GBDT меняется от  $-0.2$  до  $0.0$  в зависимости от центральности. Для получения одинаковой эффективности для трех диапазонов по центральности были выбра-

ны пороги  $0.0$  для диапазонов 0–10% и 10–30% и порог  $-0.2$  для диапазона 30–100%.

Поскольку зависимости ширины струй, заряда струй и множественности заряженных частиц в струе ведут себя по-разному в присутствии эффекта гашения струй и без него, то стоит ожидать, что применение дерева решений, натренированного на данных без гашения струй, на данных с гашением струй приведет к изменению соотношения эффективностей выделения кварковых и глюонных струй.

#### 5. ИЗМЕРЕНИЕ ДОЛИ (ФРАКЦИИ) ГЛЮОННЫХ СТРУЙ

Было проведено сравнение фракции глюонных (или кварковых) струй измеренной с помощью обученной модели GBDT с истинной фракцией, известной из моделирования. Данные фракции нельзя сравнивать напрямую, без учета эффективности GBDT при идентификации кварковых и глюонных струй. Если  $N$  — полное число событий,  $p$  — истинная доля глюонов,  $q$  — кварков,  $e_p$  и  $e_q$  — соответствующая эффективность для глюонов и кварков, то измеренная фракция  $\pi$  глюонных струй:

$$\pi = \frac{Npe_p + Nqe_q}{N} = pe_p + qe_q. \quad (3)$$

При условии, что  $p + q = 1$ , мы можем выразить  $p$  как

$$p = \frac{\pi - e_q}{e_p - e_q}. \quad (4)$$

Из (4) видно, почему нужно оптимизировать порог по GBDT именно по разности эффективностей. Чтобы измерение имело смысл, должно быть  $\pi > e_q$ , т.е. эффективность к кварковым струям не должна быть слишком высокой. Также  $e_p > e_q$ . Учитывая, что  $p < 1$ , условие можно записать в виде  $e_q < \pi < e_p$ . Из 3 видно, что эти условия выполняются при  $e_p < e_q$ .

Статистические ошибки для  $\pi$ ,  $e_p$ ,  $e_q$  вычисляются по формуле для мультиполиномиального распределения. Ошибка  $\delta_p$  определяется по формуле:

$$\frac{\delta_p}{p} = \sqrt{\frac{\delta_\pi^2 + \delta_{e_q}^2}{(\pi - e_q)^2} + \frac{\delta_{e_p}^2 + \delta_{e_q}^2}{(e_p - e_q)^2}}. \quad (5)$$

На рис. 7 приведено итоговое значение для измеренной на независимой статистике Монте-Карло фракции глюонных струй для разных диапазонов по центральности. Истинное значение для доли глюонных струй при этом равно  $0.2$ . Статистические ошибки находятся в пределах 5%. Результаты на данных с гашением струй и без него совпадают.

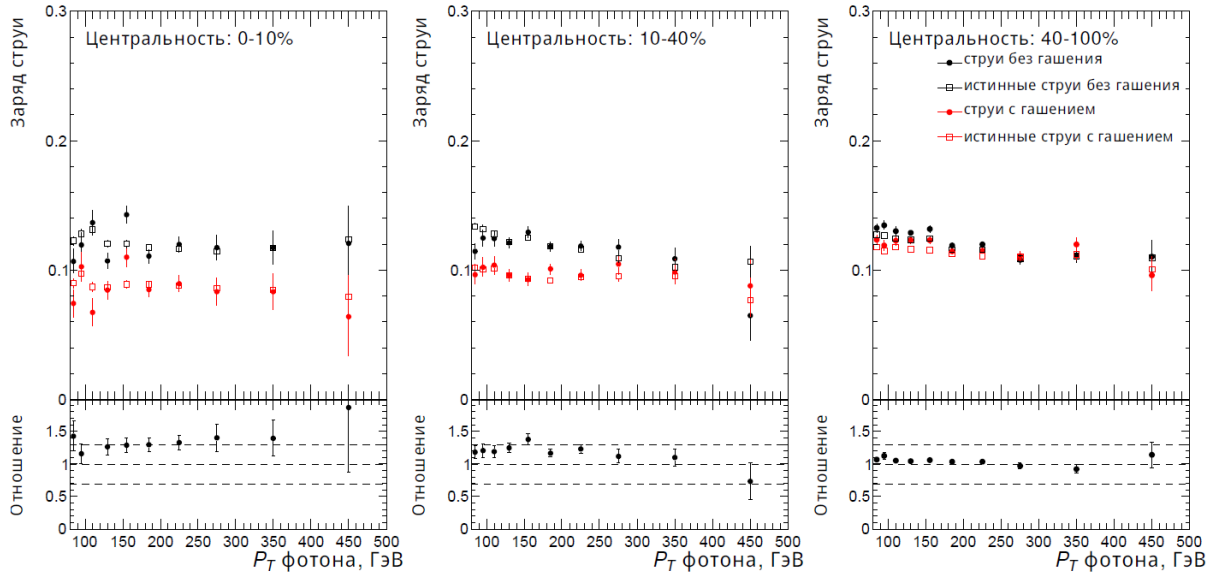


Рис. 4. Измеренные значения переменных отдельно для кварковых и глюонных струй в зависимости от центральности соударений. Приведены зависимости для вариантов с гашением и без гашения сигнальной струи

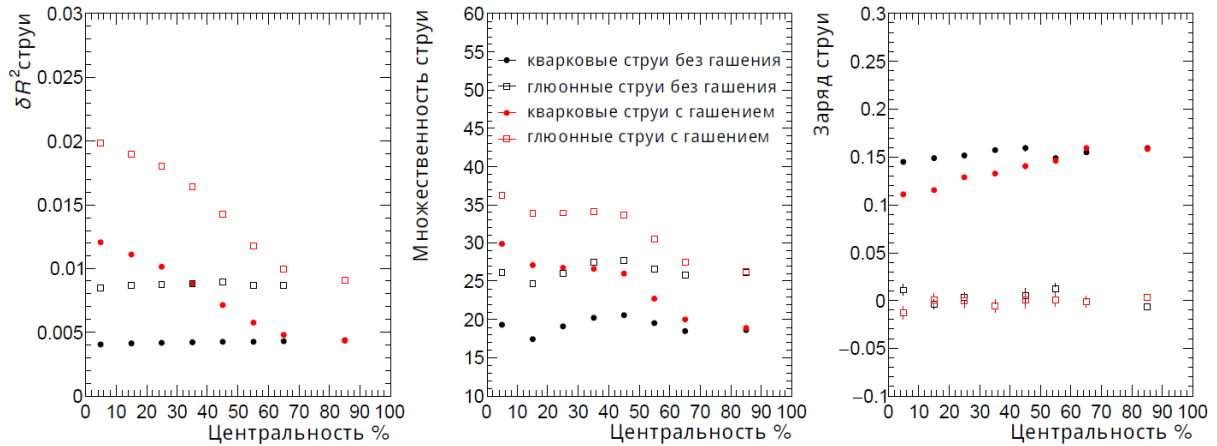


Рис. 5. Измеренные значения переменных отдельно для кварковых и глюонных струй в зависимости от центральности соударений. Приведены зависимости для вариантов с гашением и без гашения сигнальной струи

## 6. ПРИМЕНЕНИЕ МЕТОДИКИ В ЭКСПЕРИМЕНТЕ

Возникает вопрос: как применять разработанную методику в реальном эксперименте. Применение модели, обученной на Монте-Карло данных с эффектом гашения струй будет сильно зависеть от того, насколько смоделированные данные совпадают с экспериментальными. В настоящее время Монте-Карло модели не полностью описывают экспериментальные данные. Однако это оспаривается рабочим вариантом, так как можно пробовать по-разному настроить Монте-Карло генераторы PYQUEN и HYDJET. В случае обучения на Монте-Карло данных без эффекта гашения струй (которые описывают экспериментальные данные pp-соударений значительно лучше) возникает другая проблема, связанная с изменением формы струй в экспери-

ментальных данных в результате наличия эффекта гашения струй.

Решением в данном случае является дополнительная коррекция переменных в экспериментальных данных с целью удаления эффекта гашения струй перед поступлением на вход обученной модели. Значения переменных  $V_q$  (множественность, ширина или заряд) для каждой струи поправляются на фактор, полученный из разницы средних значений переменных с эффектом гашения и средних значений переменных без эффекта гашения:

$$V_{гаш}^{корр} = V_{гаш} - (\langle V_{гаш} \rangle - \langle V \rangle). \quad (6)$$

При этом приходится предполагать, что фактор одинаков для глюонных и кварковых струй, так как в эксперименте мы видим смесь струй в неизвестной пропорции. Из рис. 5 видно, что это предположение разумно. Коррекционный фактор вычис-



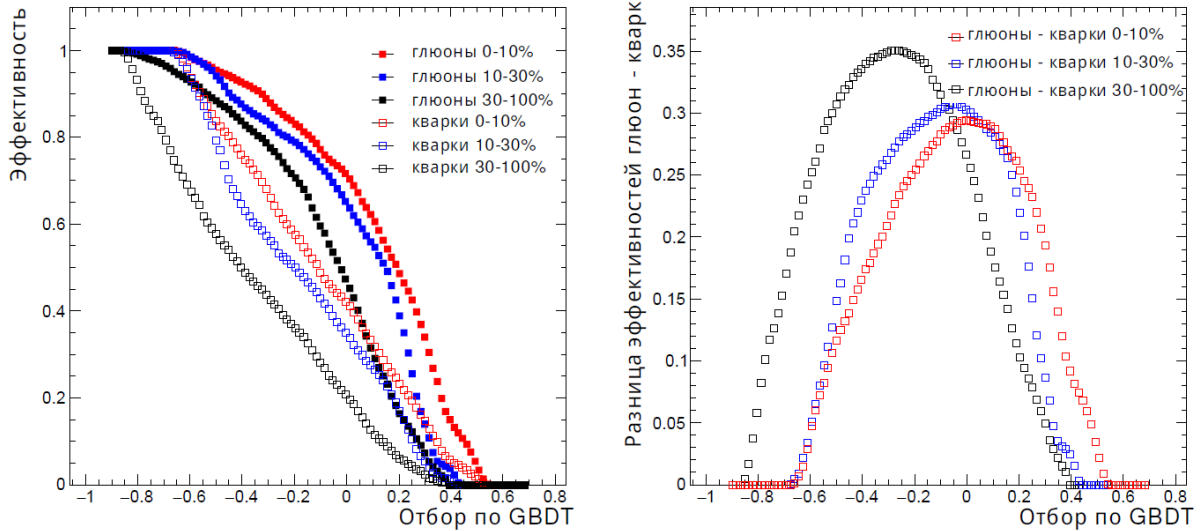


Рис. 6. Зависимость эффективности идентификации глюонных и кварковых струй (слева) и их разности (справа) в зависимости от порога по выходному значению GBDT для трех интервалов по центральности. Тренировка модели проведена на данных без эффекта гашения струй

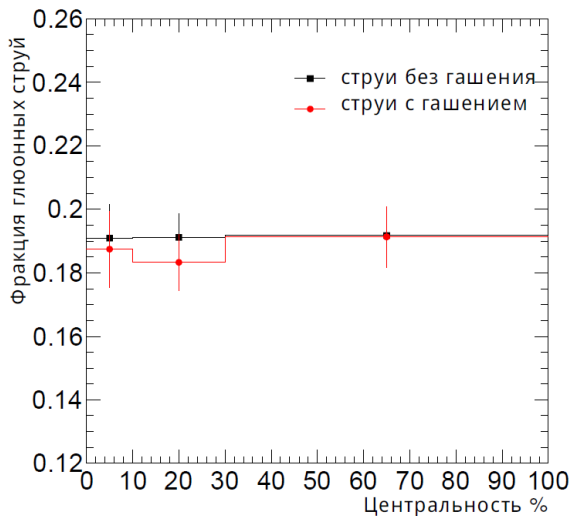


Рис. 7. Измеренная фракция глюонных струй в зависимости от центральности для данных без эффекта гашения струй и с эффектом гашения струй. Истинная фракция, заложенная в Монте-Карло, равна 0.2

ляется в зависимости от центральности соударения и аппроксимируется полиномом третьей степени. В результате получают откорректированные струи, значения переменных которых близки к струям без эффекта гашения струй. Они поступают на вход модели машинного обучения, натренированной на данных без эффекта гашения струй.

Для проверки применимости описанной методики используется набор Монте-Карло данных с эффектом гашения струй (имитирующий экспериментальные данные) с откорректированными переменными. Переменные поступают на вход модели машинного обучения, натренированной на данных без эффекта гашения струй. При пороговых значениях

GBDT 0.05, 0.0, -0.05 (для соответствующих диапазонов по центральности 0–10%, 10–30%, 30–100%) фракция глюонных струй варьируется в пределах от 0.17 до 0.22, что близко к значениям, показанным на рис. 7. Для конечной оптимизации порогов требуется обучение на большей статистике. Также следует более детально изучить зависимость результата от выбранного порога.

## ЗАКЛЮЧЕНИЕ

Исследованы переменные, характеризующие форму, множественность и заряд струи при взаимодействии тяжелых ионов. Разработана процедура вычитания фона, позволяющая учесть влияние тяжелоионного события на форму струи. Показано, что в случае возникновения кварк-глюонной плазмы множественность и ширина струи начинают сильно зависеть от центральности события, в то время как заряд струи остается постоянным. Это позволяет использовать данные переменные для регистрации возникновения кварк-глюонной плазмы, а также исследовать характеристики КПП путем сравнения с разными моделями. Данные переменные принимают разные значения для кварковых и глюонных струй для всего диапазона по центральности соударений, что позволяет их использовать в машинном обучении для идентификации кварковых и глюонных струй. Полученные с помощью машинного обучения значения доли глюонных струй совпадают с истинными значениями из Монте-Карло как в случае с наличием эффекта гашения струй, так и без него. Также показана возможность коррекции значений переменных при наличии эффекта гашения струй для применения на модели, обученной без эффекта гашения струй, что может оказаться важным при анализе экспериментальных данных.

- [1] Шуряк Э.В. // ЯФ. **28**. 796. (1978).
- [2] Baier R., Schiff D., Zakharov B.G. // *Annual Review of Nuclear and Particle Science*. **50**:37-69 (2000).
- [3] Gras P., Höche S., Kar D. et al. // *Journal of High Energy Physics*. **2017** Art. number: 91 (2017).
- [4] Gyulassy M., Plumer M. // *Phys. Lett.* **B243**. 432. (1990).
- [5] Gyulassy M., Wang X. // *Nucl. Phys. B*. **420**. 583. (1994).
- [6] Wang X., Gyulassy M. // *Phys. Rev. Lett.* **68**. 1480. (1992).
- [7] Розенталь И.Л., Смигирев А.М. // Физика элементарных частиц и атомного ядра. **34**, Вып. 1. 165. (2003).
- [8] Dissertori G., Knowles I. G., Schmelling M. 12 Differences Between Quark And Gluon Jets. In: *Quantum Chromodynamics: High Energy Experiments and Theory. International Series of Monographs on Physics*. p. 377. (Oxford, 2009; online edn, Oxford Academic, 1 Jan. 2010)
- [9] Casalderey-Solana J., Gulhan D.C., Milhano J.G. et al. // *J. High Energy Phys.* **19**. (2014).
- [10] CMS and ATLAS collaborations. // «Distinguishing quark and gluon jets at the LHC». Parton Radiation and Fragmentation from LHC to FCC-ee : CERN, Geneva, Switzerland, November 22-23, 2016, 73-78. Contribution to: eeJets2016, pp. 73-78.
- [11] Lokhtin I.P., Snigirev A.M. // *Eur. Phys. J. C*. **45**. 211. (2006).
- [12] CMS Collaboration // *JINST* **3**. S08004. (2008).
- [13] Cacciari M., Salam G.P., Soyez G. // *JHEP* **4** 063. (2008). arXiv:0802.1189.
- [14] Cacciari M., Salam G.P., Soyez G. // *Eur. Phys. J. C*. **72**. 1896 (2012).
- [15] Kodolova O., Vardanian I., Nikitenko A., Oulianov A. // *Eur. Phys. J. C*. **50** 117. (2007).
- [16] CMS collaboration // «Quark-gluon Jet Discrimination At CMS», arXiv:1409.3072v1
- [17] Chatrchyan S. et al. (CMS Collaboration) // *Phys. Rev. C*. **84**, 024906 (2011).
- [18] Hoecker A., Speckmayer P., Stelzer J. et al. // «TMVA: Toolkit for Multivariate Data Analysis». PoS A CAT 040 (2007) [physics/0703039].
- [19] Образцов С.В., Кодолова О.Л., Варданян И.Н. и др. // *Вестн. Моск. ун-та. Физ. Астрон.* № 3. 9. (2018). (O.L. Kodolova, S.V. Obraztsov, I.N. Vardanyan et al. // *Moscow Univ. Phys. Bull.* **73**, N 3. 241. (2018))

## Separation of quark and gluon jets in heavy ion collisions at high energies

S. V. Obraztsov<sup>a</sup>, O. L. Kodolova<sup>b</sup>

*Lomonosov Moscow State University. Skobeltsyn Institute of Nuclear Physics (SINP MSU)  
Moscow 119991, Russia*

*E-mail: <sup>a</sup>stepan.obraztsov@cern.ch, <sup>b</sup>olga.kodolova@cern.ch*

The possibility of separating quark and gluon jets in heavy ion collisions, both in the presence of the jet quenching effect and in its absence, has been investigated. Using machine learning methods, a technique has been developed to recover the fraction of quark and gluon jets, using variables characterizing the shape, multiplicity, and charge of the jet. The potential of using these variables to study the characteristics of the dense medium that arises during high-energy ion collisions has been demonstrated.

PACS: 13.85.-t

*Keywords:* heavy ion physics, hadronic jets, machine learning.

*Received 28 May 2023.*

English version: *Moscow University Physics Bulletin*. 2023. **78**, No. 5, pp. 634–640.

### Сведения об авторах

- Образцов Степан Владимирович — мл. науч. сотрудник; e-mail: [stepan.obraztsov@cern.ch](mailto:stepan.obraztsov@cern.ch).
- Кодолова Ольга Леонидовна — канд. физ.-мат. наук, ст. науч. сотрудник; e-mail: [olga.kodolova@cern.ch](mailto:olga.kodolova@cern.ch).