

Московский ордена Трудового Красного Знамени
физико-технический институт
(государственный университет)
Факультет общей и прикладной физики
Кафедра фундаментальных и прикладных проблем физики
микромира
Объединенный институт ядерных исследований
Учебно-научный центр

Кожевников Д.А.

Исследование двухструйных событий в эксперименте ATLAS.

Бакалаврская работа

*Научный руководитель
к.ф.-м.н Бойко И.Р.*

Дубна Июнь 2011

Содержание

1	Введение	3
2	Установка ATLAS	5
3	Распознавание струй, образующихся в результате адронизации красивых кварков	7
4	Метод выполнения	9
5	Результаты работы	10
5.1	Анализ данных Монте-Карло моделирования	10
5.2	Анализ двухструйных событий	12

1 Введение

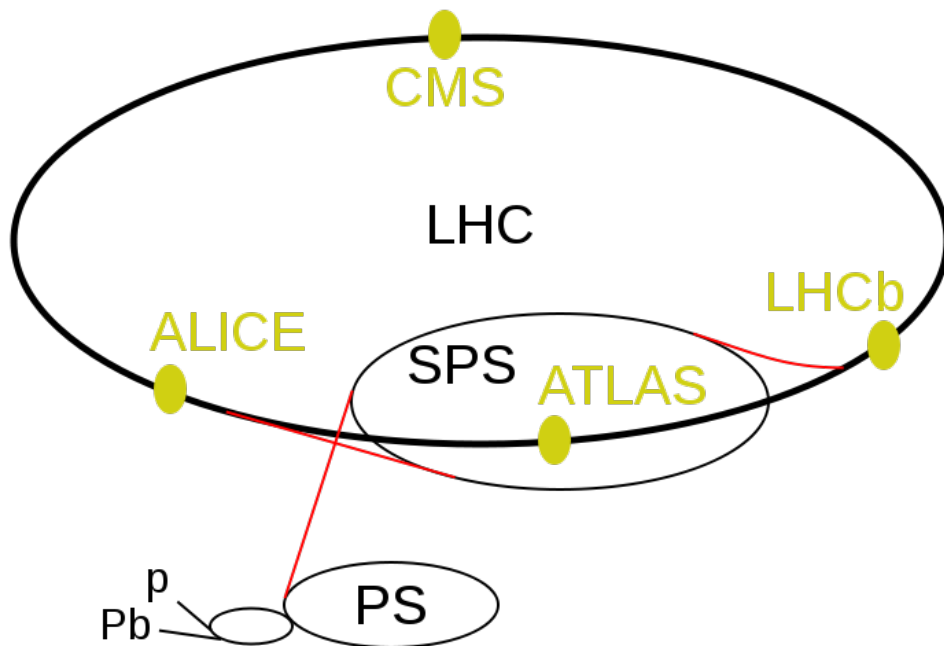


Рис. 1: Детекторы и предускорители БАК. Траектория протонов p (и тяжёлых ионов свинца Pb) начинается в линейных ускорителях (в точках p и Pb , соответственно). Затем частицы попадают в бустер протонного синхротрона (PS), через него — в протонный суперсинхротрон (SPS) и, наконец, непосредственно в туннель БАК.

ATLAS (A Toroidal LHC ApparatuS) - один из четырёх основных экспериментов на Большом адронном коллайдере (LHC) в Европейском центре ядерных исследований (CERN) в Женеве. С апреля 2010 года ускоритель работает на столкновение протонов с суммарной энергией 7 ТэВ. Одноименный детектор ATLAS разрабатывался как многоцелевой. Когда создаваемые Большим Адронным Коллайдером встречные протонные пучки сталкиваются в центре детектора, возникают различные частицы с широким спектром энергий. Вместо того, чтобы сосредотачиваться на специфическом физическом процессе, ATLAS разработан для измерения самого широкого диапазона сигналов от рождения и распада частиц. Это гарантирует, что, независимо от формы, которую могли бы принять любые новые физические процессы или частицы, ATLAS будет в состоянии обнаружить их и измерить их свойства.

Основными пунктами программы эксперимента являются:

- Поиск бозона Хигса - недостающей частицы Стандартной модели. Если бозон Хигса не будет обнаружен, то ожидается, что будут обнаружены другие механизмы нарушения электрослабой симметрии (такие как technicolor), объясняющие те же самые явления.
- Измерение свойств красивых частиц для подтверждения или опровержения стан-

дартной модели, исследования нарушения CP-инвариантности.

- Изучение свойств t -кварка и взаимодействия его с другими частицами, что может предоставить косвенную информацию о деталях Стандартной модели.
- Исследование новой физики, в частности связанной с теорией суперструн, различные варианты которой предполагают существование новых очень массивных частиц, которые во многих случаях распадаются в высокоэнергетический кварк и тяжёлые стабильные частицы.

Поскольку многие из этих физических процессов сопровождаются рождением b -кварков, в эксперименте ATLAS важно качественно распознавать струи, образовавшиеся в результате адронизации b -кварков. В связи с этим имеет практическое значение исследование свойств алгоритмов распознавания струй от b -кварков. Целью данной работы является исследование свойств алгоритмов распознавания b -струй, определение эффективности отбора b -струй и эффективность подавления легких струй. Сопоставление полученных характеристик на основе Монте-Карло моделирования с результатами применения тэгов для анализа двухструйных событий.

2 Установка ATLAS

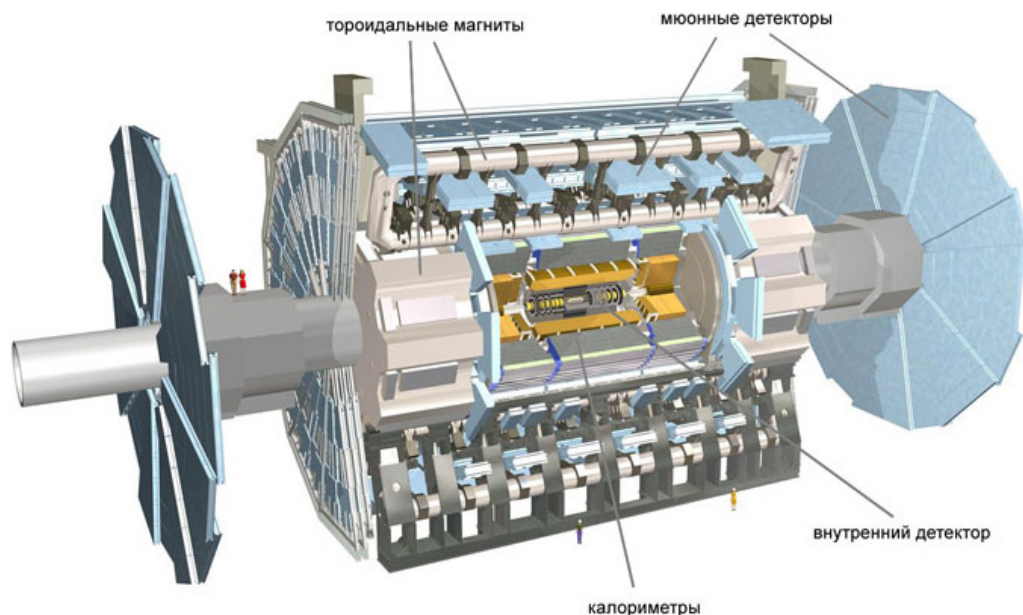


Рис. 2: Общий вид детектора ATLAS в разрезе и его основные компоненты. Рис. с сайта atlas.ch

Детектор ATLAS [1] представляет из себя классический многоцелевой детектор в коллайдерном эксперименте. Состоит из внутреннего детектора, окруженного калориметрами, а снаружи расположены мюонные детекторы. Длина детектора - 44 метра, диаметр - 25 метра. Внутренний детектор состоит из детекторов для идентификации частиц и точного измерения их траекторий. Ближе всего к оси пучка расположен пиксельный детектор, затем идет полупроводниковый трековый детектор. Третий компонент внутреннего детектора - трековый детектор переходного излучения. Внутренний детектор помещен в магнитное поле. Внутренний детектор окружен калориметрической системой. Электромагнитный калориметр состоит из пластин металла-поглотителя, в зазорах между которыми находится жидкий аргон, служащий чувствительным материалом. Адронный калориметр, следует за электромагнитным и состоит из чередующихся пластин поглотителя и пластин органического сцинтиллятора. Мюонный спектрометр состоит из дрейфовых трубок. Наружная магнитная система создает неоднородное магнитное поле со средней индукцией около 0.5 тесла.

Объем данных, поступающих с детектора, велик, и для их обработки используется система распределенных вычислений GRID. Она имеет несколько уровней организации. Базовыми элементами являются компьютерные узлы Tier0, Tier1, Tier2. Центром является Tier0, расположенный в ЦЕРНе. Он осуществляет первичную обработку данных, их калибровку, создает файлы информации для событий. После первичной обработки эти файлы рассылаются в 10 вычислительных центров Tier1, где проводится основной массив вычислений по реконструкции данных измерений. Эти результаты в форме файлов с реконструированными данными рассылаются в Tier2, где становятся доступными физикам, выполняющих анализ. Для индивидуальной работы или работы небольшой группы физиков по определенной теме предназначены локальные вычислительные базы

Tier3, функции которой уже может выполнять отдельный персональный компьютер.

При столкновении протонов при энергиях LHC, в результате их взаимодействия, образуются адронные струи.

Адронная струя — это совокупность адронов летящих в одном направлении. Средний угол раствора струи уменьшается с ростом ее энергии, т.е. адроны все сильнее группируются вокруг направления вылета кварка. При энергии 20 ГэВ составляющие струю адроны занимают 5% от полного телестного угла.

В эксперименте ATLAS принято, что минимальная энергия струи 7 ГэВ, а составляющие её адроны летят в пределах угла $\Delta R = \sqrt{\phi^2 + \eta^2}$ Реконструкция струй использует высокую степень структурированности калориметров ATLAS. Калориметры разбиты на ячейки с размерами по углам $\Delta\eta * \Delta\phi = 0.1 * 0.1$. Кроме того, для алгоритмов реконструкции важно высокое разрешение трековой системы. Восстановление вершины соударения происходит с точностью 150-200 мкм.

Есть несколько типов данных, используемых в ATLAS. "Сырые" данные - это те, что напрямую снимаются с детектора. Далее данные реконструируются и складываются в формате ESD (Event summary data). После, данные фильтруются и складываются в удобном, для последующего анализа виде в AOD (Analysis Object Data).

3 Распознавание струй, образующихся в результате адронизации красивых кварков

Рассмотрим событие в ЛHC. При столкновении протонов интенсивно взаимодействуют только два партона (жесткое рассеяние), который адронизуются в струи частиц с большим поперечным импульсом. Другие партоны адронизуются в струи частиц с малым поперечным импульсом и вылетают из детектируемого объема. Частицы в струях от от провзаимодействовавших партонов распадаются на (квази)стабильные частицы, которые уже детектируются. Большинство событий - "мягкие с низким поперечным импульсом (события *minimum bias*), однако практически вся интересная физика приходится на "жесткие" события.

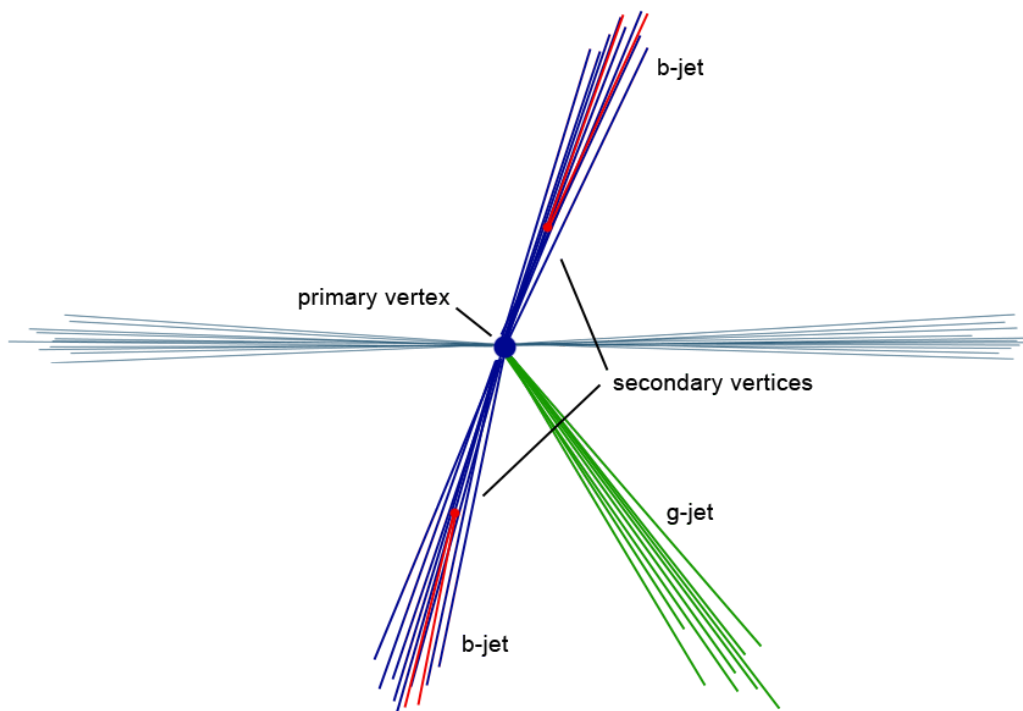


Рис. 3: Конфигурация события в ЛHC

В первичной вершине происходит вся адронизация, затем некоторые частицы, тау-лептоны и *b*-, *c*-адроны, распадаются образуя вторичные вершины. Они могут пролетать значительные расстояния, которые могут быть измерены в детекторе ATLAS. Благодаря этому можно идентифицировать *b*, *c*, *tau* частицы. Струи, образующиеся в результате адронизации легких кварков будем называть "легкими а струи от *b*-кварков - *b*-струями.

B-струи обладают рядом характерных свойств, которые могут быть использованы для их отделения от легких струй. Наиболее важным свойством является относительно большое время жизни *B*-адронов, около 1.5 пс. Благодаря этому, *B*-адроны имеют измеримые длины полета в несколько миллиметров. Распад *B*-адрона во вторичной вершине может быть идентифицирован путем измерения прицельных параметров (IP) треков,

идущих из вторичной вершины.

- $IP > 0$, если точка максимального сближения лежит за вершиной по направлению струи.
- $IP > 0$, если точка максимального сближения лежит перед вершиной по направлению струи.

Все алгоритмы опираются на отношение правдоподобия [2], для создания переменной, которая называется весом струи (weight), для разделения b-струй от других. Все струи, имеющие вес больше определенного (W_{cut}^b) определяются как b-струи. Струи с весом меньше чем $W_{cut}^l < W_{cut}^b$. При одних и тех же ограничениях на вес струи результаты работы тагеров различаются для разных импульсов струй.

- JetProb - использует отрицательную часть распределения значимости поперечного IP, полученное из расчета вероятности совместимости треков с первичной вершиной.
- IP2D - более сложный алгоритм, использующий распределение значимости IP, рассчитанное в поперечной плоскости и для IP3D кроме того в продольной проекции.
- SV1/2 - использует вторичные вершины и строит вес струи опираясь на несколько одного или более-мерного распределения переменной как, например, вершина массы.
- IP3DSV1 основан на сумме весов тагеров IP3D и SV1 и на данный момент является основным.

Есть и другие алгоритмы, в частности основывающиеся на наличии в струе мягких лептонов.

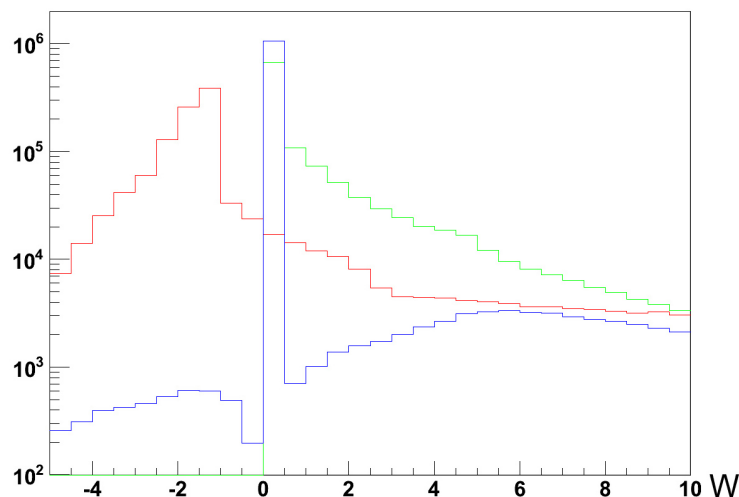


Рис. 4: Распределение струй по весу для тагеров IP3DSV1 - красным, JetProb - зеленым и SV0 - синим.

4 Метод выполнения

Для первичной обработки данных использовалась ATHENA [3] analysis framework разработанная группой ATLAS. Athena состоит из нескольких пакетов, которые позволяют выполнять следующие задания:

- Преобразование Данных
- Анализ данных
- Монте-Карло моделирование

Доступ к данным осуществляется через запрос к системе GRID [4]. Программный код, выполняющий анализ, поступает в систему GRID и находит необходимые файлы с экспериментальными данными, запрашиваемые программой физического анализа. Результаты работы программы анализа возвращаются для дальнейшей обработки. Из реконструированных данных представленных в AOD с помощью ATHENA и PANDA, необходимая информация о событиях сохранялась в формате Tree, как ROOT [5] файл. Окончательный анализ и представление данных осуществляется с помощью системы обработки данных физики высоких энергий ROOT.

Для реальных данных производился дополнительный отбор двухструйных событий, с противоположно направленными струями. Продолжение одной из струй и другая струя должны лежать в пределах конуса с $\Delta R < 0.4$

Для каждого события из специальных контейнеров данных в AOD-файле сохранялась следующая информация: 1) Из контейнера данных "AntiKt4H1TopoJets": число реконструированных струй в событии, лоренц-вектор импульса, энергия струи и вес струи, присвоенный тагером IP3DSV1, JetProb и SV0. 2) Из контейнера "Staco": число реконструированных мюонов и электронов в событии, лоренц-вектор импульса, энергия лептона. 3) Из контейнера "SpclMC": истинная информация для Монте-Карло моделирования, - тип частицы в соответствии со стандартом PDG [6].

Результат представляет из себя коллекцию гистограмм, сложенных в ROOT файлы.

5 Результаты работы

5.1 Анализ данных Монте-Карло моделирования

В работе проводится сравнение алгоритма распознавания b -струй JetProb, SV0 и комбинации IP3D + SV1 на основе данных Монте-Карло моделирования процесса $ppb \rightarrow Zbbb$ при энергии 3.5 ТэВ на нуклон. Число событий нормировано на интегральную светимость ЛНС 10 fb⁻¹. Для удобства сравнения для тагера JetProb берется $-\log(\text{weight})$, благодаря чему распределение струй по весу для тагеров схожи (см. рисунок 3). Большинство струй имеет небольшой вес, что и следовало ожидать.

Из сравнение распределений "легких" струй и b -струй для тагера JetProb (см. рисунок 5.1) видно, что тагер начинает работать при ограничениях на вес струй начиная с $w > 6$, когда количество b -струй среди отобранных значительно превышает количество легких струй. Аналогично для тагера IP3DSV1 получаем, что приемлемое ограничение на вес струи примерно 3 (см. рисунок 5.1).

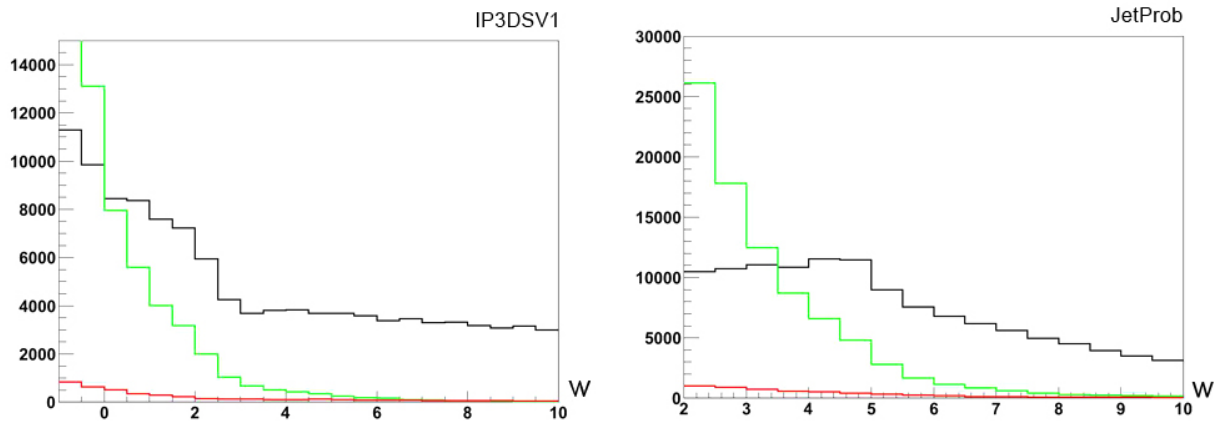


Рис. 5: Слева: распределение струй по весу для тагера IP3DSV1. Черным показано распределение b -струй, красным - c -струй и зеленым для легких струй. Справа: распределение струй по весу для тагера JetProb. Черным показано распределение b -струй, красным - c -струй и зеленым для легких струй.

Для определения свойств тагеров выбраны следующие характеристики:

- Эффективность отбора b -струй: количество отобранных b -струй при ограничении на вес по отношению к полному количеству b -струй.
- Чистота отобранных b -струй: количество отобранных b -струй при ограничении на вес по отношению к полному количеству отобранных струй.
- Эффективность подавления легких струй: аналогично эффективности отбора b -струй.

Эти характеристики могут быть получены только на основе Монте-Карло моделирования, когда известна достоверная информация о струе. Сначала подсчитывается

количество "легких" и b -струй (из Монте-Карло известно, каким кварком порождена струя), затем подсчитывается число струй, оставшихся после наложения условия на вес струи.

Полученные характеристики отличаются для струй с разным импульсом, поэтому их зависимости от ограничения на вес струи были построены для струй разных интервалов поперечного импульса: от 7 ГэВ до 34 ГэВ через 1 ГэВ, от 34 ГэВ до 40 ГэВ, от 40 ГэВ до 50 ГэВ и от 50 ГэВ. На рисунке 5.1 приведены зависимости для двух периодов. Видно, что для струй с большим поперечным импульсом эффективность отбора b -струй лучше, чем для струй с меньшим поперечным импульсом. Однако, эффективность подавления легких струй несколько ухудшается с ростом поперечного импульса.

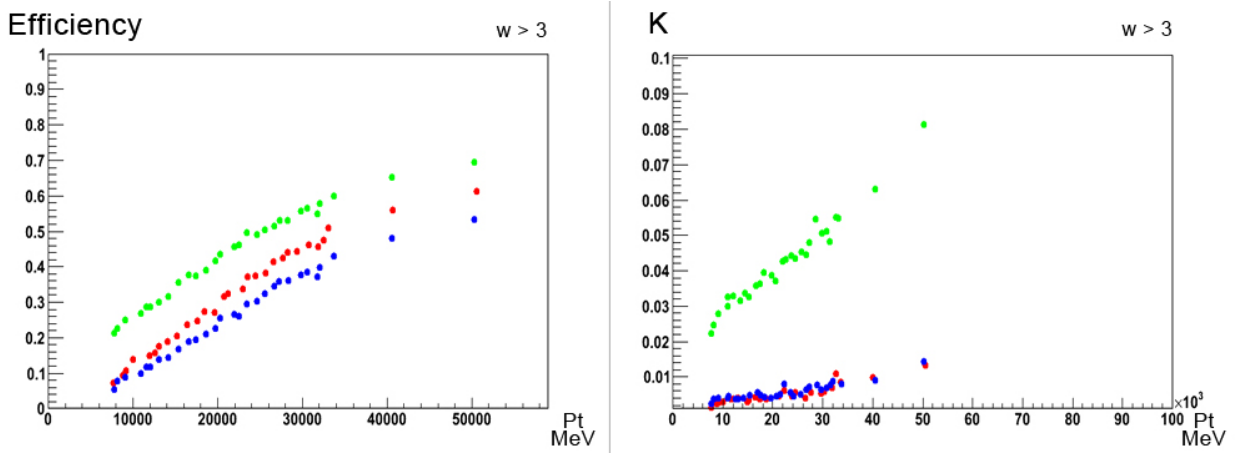


Рис. 6: Слева: эффективность отбора b -струй в зависимости от поперечного импульса струи при $w \geq 3$ для тагеров: IP3DSV1 - красным, JetProb - зеленым и SV0 - синим. Справа: эффективность подавления легких струй в зависимости от поперечного импульса струи при $w \geq 3$ для тагеров: IP3DSV1 - красным, JetProb - зеленым и SV0 - синим.

На струях с низким поперечным импульсом эффективность работы тагеров низкая, поэтому в анализе следует отбросить струи с поперечным импульсом ниже определенного значения. В качестве такового было выбрано значение $P_t = 20$ ГэВ.

Чистота отобранных струй зависит от начального соотношения легких и b -струй и эту характеристику не применить напрямую к обработке данных. Но её можно использовать для сравнения тагеров на одном и том же наборе событий. Покажем это для интервала по поперечному импульсу струи от 30 до 31 ГэВ. Для этого занесем в таблицу значения на минимальный вес струи, чистоту и коэффициент подавления легких струй для трех тагеров, при эффективности 0.42.

$Efficiency = 0.49$	$Weight$	$Purity$	$LightJetrejection$
JetProb	5.5	0.945	66
IP3DSV1	5.2	0.990	303
SV0 ($e = 0.42$)	1	0.962	101

Таким образом, при одной и той же эффективности отбора b -струй надо устанавливать разные значения на минимальный вес струи. При этом для тагера IP3DSV1 чистота

отобранных струй будет лучше, а коэффициент подавления легких струй в несколько раз выше.

5.2 Анализ двухструйных событий

В эксперименте ATLAS "сырые" данные разделяются на 4 основных физических потока данных: $e\gamma$, $\mu\mu$, JetTauEtmis и μb . Каждый поток состоит из событий в которых срабатывают один или несколько триггеров.

Для исследования алгоритмов распознавание b -струй на основе реальных данных были обработаны данные калориметрического потока отмеченные тегом "JetTauEtmis" за 2011 год, периоды $A - E$. В отличие от данных Монте-Карло моделирования, здесь неизвестно достоверно, от какого кварка произошла струя. Для того, чтобы приписать струе определенный кварк, от которого она берет начало и используются алгоритмы распознавания струй. Для проверки качества их работы Игорем Бойко и Михаилом Демичевым был предложен следующий способ, основанный на физических соображениях, что красивые кварки рождаются в основном парами, причем они массивные и напололам уносят основную энергию события:

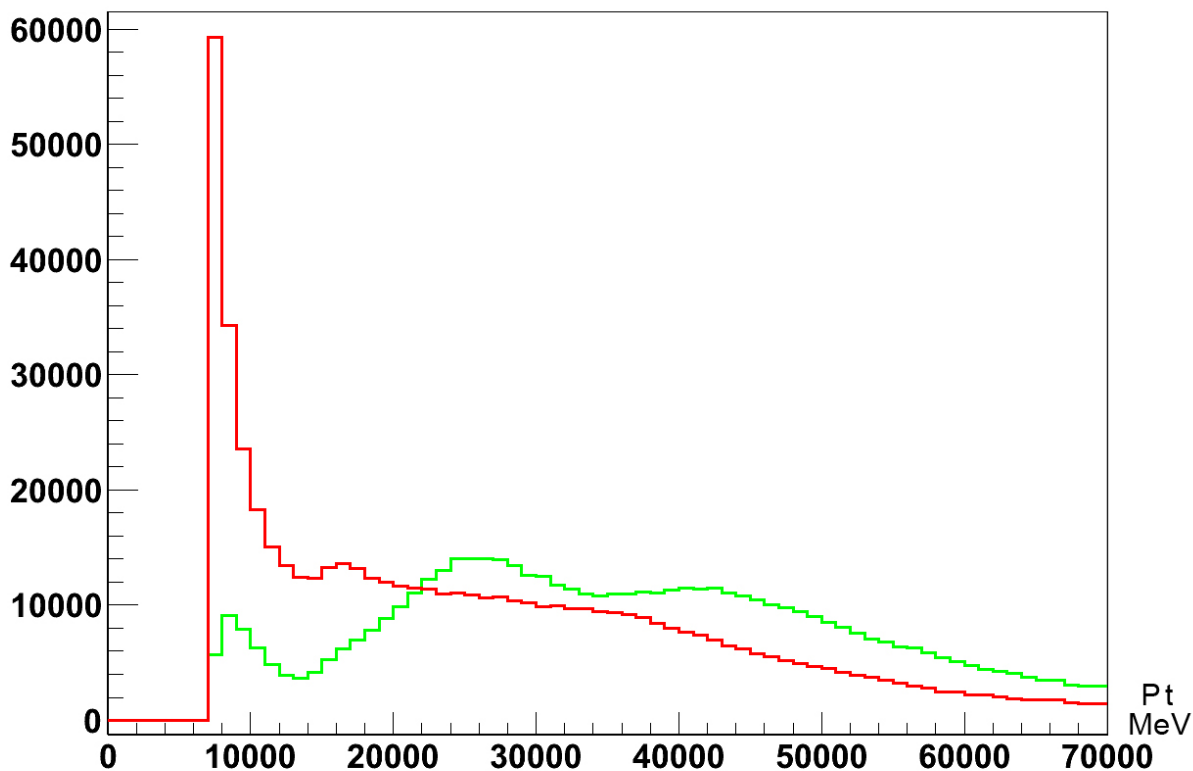


Рис. 7: Распределение струй по поперечному импульсу. Первая струя, с большим p_t - зеленым, вторая - красным.

1. Изначально из всех данных отбираются события в которых ровно две струи, при-

чем противоположно направленных, в поперечном направлении в угле $\Delta R = 0.4rad$, который совпадает с шириной самой струи.

2. Из Монте-Карло известно, что таггеры на мягких струях работают либо ненадежно, либо неэффективно, в зависимости от ограничения на минимальный вес струи. Поэтому отбрасываются события, если одна или две струи имеют $Pt > 20$.
3. Сортируем струи по поперечному импульсу, считаем струю с большим поперечным импульсом первой струю, с меньшим - второй струей 5.2. Разносим события по четырем возможным категориям: $[B, B]$, $[B, L]$, $[L, B]$, $[L, L]$ следующим образом:

Для каждой струи смотрим на её вес W . Если $W > WB = 6$, то считаем что это b-струя. Если $W < WL = 1$, считаем струю "легкой". Если вес хотя бы одной струи оказывается $WL < W < WB$, такая струя считается неопределенной и событие не учитывается.

4. По оставшимся двухструйным событиям считаем следующие отношения:

$$rL = \frac{N_{events}(W_{jet2} < W_L^{cut})}{N_{events}(W_{jet2} < W_L^{cut} \cap W_{jet1} > W_B^{cut})} = \frac{|[B, L]| + |[L, L]|}{|[B, L]|}$$

$$rB = \frac{N_{events}(W_{jet2} > W_B^{cut})}{N_{events}(W_{jet2} > W_B^{cut} \cap W_{jet1} > W_B^{cut})} = \frac{|[L, B]| + |[B, B]|}{|[B, B]|}$$

и

$$R = \frac{rL}{rB}$$

Должно быть $R \gg 1(10 - 20)$ и отношение R не должно зависеть от начального соотношения "легких" и b-струй. Для разных таггеров отношение окажется разным. Если при увеличении WB отношение стремится к значению $R < 5$, алгоритм определения b-струй на данных не работает.

Для сравнения двух таггеров при $W_l^{cut} = 1$ построена зависимость:

$$R = \frac{N_L(W_B^{cut})}{N_B(W_B^{cut})}$$

В таблице приведено сравнение зависимостей $R(w)$ для таггеров JetProb и IP3DSV1 на основе анализа двухструйных событий калориметрического потока для IP3DSV1 и $-\ln(\text{JetProb})$:

IP3DSV1	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7
JetProb	4.5	4.9	5.4	5.6	5.8	5.95	6.1	6.2	6.4	6.8	6.5	7.3	7.5

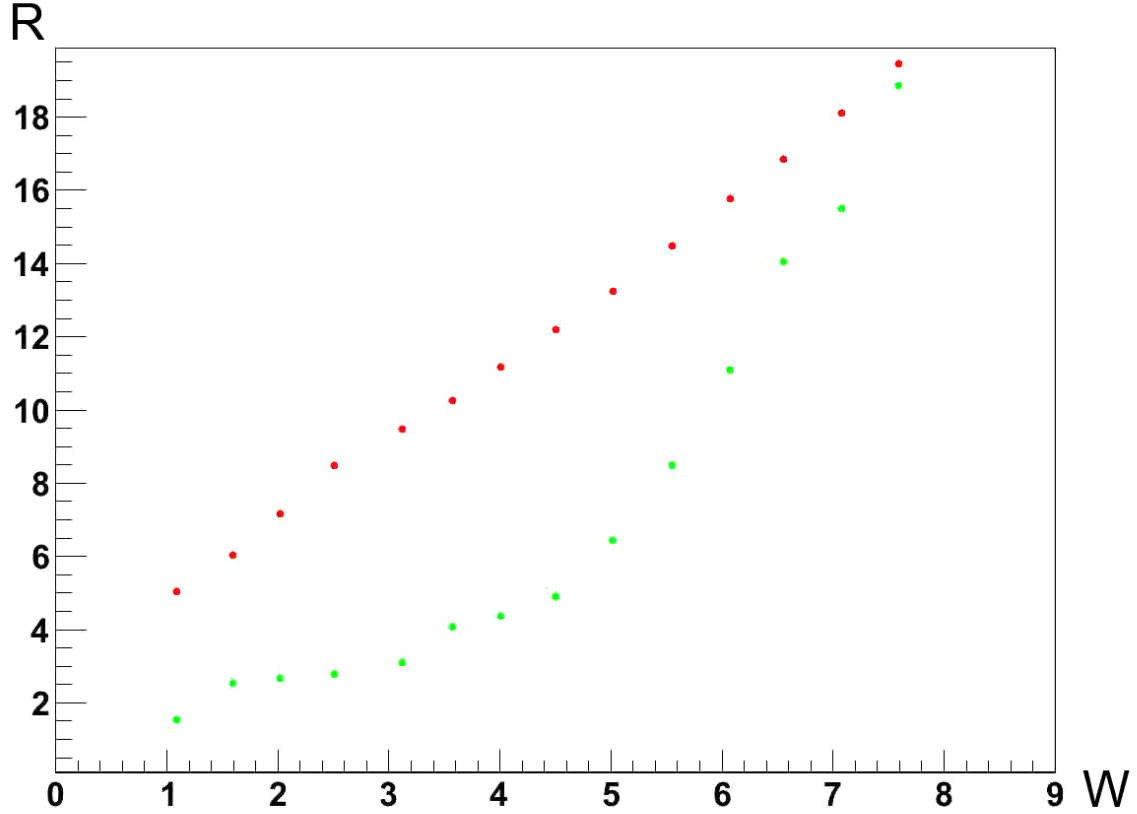


Рис. 8: Зависимость $R(W_B^{cut})$ от ограничения на вес струи для тагеров: IP3DSV1 - красным, JetProb - зеленым.

Используя характеристики тагера IP3DSV1 при ограничении на вес ($w > 3$) можно посмотреть, сколько струй будет отобрано в случае, если все струи легкие, или все струи от b -кварков. Предположим, что все струи от b -кварков, тогда условие на вес первой струи из 636409 событий оставит 337891 событий, а дополнительное условие на вторую струю 132658 событий. Аналогично, если все струи легкие из 636409 событий 51575 событий детектируются как b -струи и 4179 событий, когда обе струи от b -кварков.

Предположим, что в начальном семпле 90% - легкие струи, 10% - b -струи, тогда при наложении условия на вес первой струи останется 33790 b -струй и 46420 легких струй, которые будут тагированы как b -струи. При наложении условия на вес второй струи останется 13266 истинных b -струй и 3761 легкая струя, детектированная как b -струя. Таким образом, 78% струй из отобранных после наложения условия на вес обеих струй окажутся истинными b -струями.

Аналогично, если изначально 99% - легкие струи, и 1% - b -струи, получим что 24% струй окажутся истинными b -струями.

Для данных за периоды от А до Е из потока JetTauEtmisс за 2011 год, получено отношение

$$\frac{jet1 - b \text{ and } jet2 - b}{jet1 - b, jet2 \text{ not } b} = 0.25$$

Список литературы

- [1] The ATLAS Collaboration. *Expected Performance of the ATLAS Experiment. Volume I: Performance*. CERN-OPEN-2008-020, 2009.
- [2] Thomas Gopfer. Tagging b-jets in atlas. *ATL-PHYS-PROC-2010-014*, 2010.
- [3] The ATLAS Collaboration. *The ATLAS Computing Workbook*. 2010.
- [4] Usatlas grid.
- [5] The ROOT team. *User's Guide, 5.22*. ???, 2009.
- [6] *Review of Partical Physics, Particle Data Group. Monte Carlo Particle Numbering Scheme*. 2008.