

УДК 539.12.01

РЕШЕНИЕ ПРОБЛЕМЫ ОПТИМАЛЬНОГО ОПРЕДЕЛЕНИЯ АДРОННЫХ СТРУЙ

Ф. В. Ткачев

Институт ядерных исследований РАН, Москва

ВВЕДЕНИЕ	1177
СТАНДАРТНЫЕ АЛГОРИТМЫ ОПРЕДЕЛЕНИЯ СТРУЙ	1179
КРИТИКА «ДИНАМИЧЕСКОГО» ПОДХОДА. КИНЕМАТИЧЕСКИЕ АСПЕКТЫ	1182
ТЕОРИЯ C -НЕПРЕРЫВНЫХ НАБЛЮДАЕМЫХ	1184
КВАНТОВО-ПОЛЕВЫЕ АСПЕКТЫ. КВАРК-АДРОННАЯ ДУАЛЬНОСТЬ	1190
СТАТИСТИЧЕСКИЕ АСПЕКТЫ	1192
ОПТИМАЛЬНОЕ ОПРЕДЕЛЕНИЕ СТРУЙ	1194
ЗАМЕЧАНИЯ ОБ АЛГОРИТМИЧЕСКОЙ РЕАЛИЗАЦИИ	1198
НОВЫЕ ВОЗМОЖНОСТИ	1199
ЗАКЛЮЧЕНИЕ	1201
СПИСОК ЛИТЕРАТУРЫ	1201

УДК 539.12.01

РЕШЕНИЕ ПРОБЛЕМЫ ОПТИМАЛЬНОГО ОПРЕДЕЛЕНИЯ АДРОННЫХ СТРУЙ

Ф. В. Ткачев

Институт ядерных исследований РАН, Москва

Дан обзор систематической теории, позволившей решить проблему построения «идеального» алгоритма определения адронных струй. Теория не содержит произвольных предположений и полностью основана на положениях математической статистики и квантовой теории поля. Полученное «оптимальное» определение струй допускает эффективную алгоритмическую реализацию. Теория также содержит ряд важных опций для повышения качества обработки экспериментальных данных в случаях прецизионных измерений или малого отношения сигнал/шум.

Reviewed is a systematic theory that solves the problem of construction of an «ideal» definition of hadron jets. The theory does not contain arbitrary assumptions and is fully based on the first principles of mathematical statistics and quantum field theory. The obtained «optimal» jet definition allows an efficient algorithmic implementation. The theory also contains a number of new options for enhancing the quality of experimental data processing in cases of precision measurements or a low signal/background ratio.

ВВЕДЕНИЕ

Проблема оптимального определения адронных струй в физике высоких энергий оставалась открытой с середины 70-х гг. Теория, развитая в работах [1–6], дала систематическое и полное ее решение. Найденное определение допускает эффективную алгоритмическую реализацию и не обладает недостатками обычно используемых эмпирических алгоритмов. Кроме того, открываются новые возможности построения усовершенствованных процедур обработки данных в экспериментах на современных ускорителях элементарных частиц.

На качественном уровне адронная струя (в дальнейшем просто «струя») — это достаточно узкий пучок адронов, несущий довольно большую долю энергии события. Тот факт, что в столкновениях при высоких энергиях адронная материя имеет тенденцию рождаться в виде струй, связан с кварк-глюонной структурой адронов. Действительно, в соответствии с общепринятыми представлениями процесс столкновения при высоких энергиях происходит в два этапа. Сначала происходит жесткое столкновение квантов, из которых состоит начальное состояние, — электронов, фотонов, партонов (т. е. кварков

и глюонов) и т. п., в том числе входящих в состав сталкивающихся адронов, в результате которого рождается небольшое число энергичных квантов. Если в числе последних имеются партоны, то они в свободном состоянии существовать не могут и претерпевают превращение в наблюдаемые адроны в процессе так называемой адронизации. Если первоначально родившиеся партоны были достаточно энергичными, то естественно ожидать, что структура потока энергии на уровне наблюдаемых адронов сохранит основные качественные черты потока энергии на уровне партонов и каждому партону будет соответствовать относительно узкий пучок адронов, несущий примерно ту же энергию. Такие рассуждения дали возможность на рубеже 60–70-х гг. еще в рамках партонной модели высказать гипотезу о преобладании в достаточно жестких соударениях событий, состоящих преимущественно из адронных струй [7]. Экспериментально струи были обнаружены в процессе аннигиляции электрон-позитронной пары [8], причем было найдено, что большинство событий состоит из пары противоположно направленных струй (в системе центра масс электрон-позитронной пары), а угловое распределение струй соответствует угловому распределению пары фермион–антифермион со спином $1/2$.

После появления квантовой хромодинамики (КХД) была высказана гипотеза, что предсказания КХД для наблюдаемых типа угловых распределений кварков и глюонов можно принимать всерьез, если интерпретировать их в терминах струй, так что достаточно энергичному глюону, испущенному в поперечном направлении из кварк-антикварковой пары, также должна соответствовать струя [9]. Эти качественные соображения были существенно уточнены в работе [10], где было указано ключевое условие инфракрасной (ИК) конечности, которому должны удовлетворять наблюдаемые, чтобы их расчеты по теории возмущений в терминах партонов можно было сопоставлять с соответствующими измерениями в терминах адронных струй. ИК-конечность означает, что наблюдаемая не должна менять свое значение на состоянии, получающемся из другого заменой одной частицы на две или больше частиц, летящих в том же направлении и несущих ту же суммарную энергию.

Однако в работе [10] был рассмотрен лишь класс ИК-конечных наблюдаемых, которые строятся посредством так называемых алгоритмов определения струй (в дальнейшем — АОС). Такой алгоритм определяет, из скольких струй состоит событие, а также характеристики струй (энергия и импульс). С физической точки зрения струя — это видимый след кварка или глюона, в какой-то мере аналогичный наблюдаемому треку от, скажем, электрона. Поэтому «глядя» на струи (т. е. измеряя их угловые распределения и т. п.), мы фактически наблюдаем соответствующие партоны. Поэтому можно построить наблюдаемую, чувствительную к какой-то детали теории, рассматривая лишь партоны в низшем порядке теории возмущений, а затем подставить вместо партонов струи, чтобы получить наблюдаемую, определенную на уровне адронов и

пригодную для обработки экспериментальных данных. Если требуется учесть высшие поправки теории возмущений, то следует применять алгоритм определения струй также и к партонному состоянию в теоретических расчетах. Разумеется, в теоретических расчетах и при обработке экспериментальных данных следует применять один и тот же алгоритм определения струй.

Простота и наглядность такого подхода сделала его мощным средством физического анализа (ср. обзор [11]). При этом закрепились динамическая точка зрения на АОС как на механизм обращения адронизации. Проблемы выявились позже, когда актуальной стала задача систематического учета высших теоретических поправок в контексте более точных экспериментов. Динамическая интерпретация АОС оказалась здесь препятствием. Можно сказать, что содержание теории [1–6] состоит в указании «правильных» путей обхода этого препятствия.

1. СТАНДАРТНЫЕ АЛГОРИТМЫ ОПРЕДЕЛЕНИЯ СТРУЙ

Итак, АОС является преобразованием события P вида

$$P = \{p_1, p_2, \dots, p_{N_{\text{part}}}\} \xrightarrow{\text{АОС}} Q = \{Q_1, Q_2, \dots, Q_{N_{\text{jets}}}\}, \quad (1)$$

где p_a — 4-импульсы частиц в конечном состоянии; Q_j — 4-импульсы струй.

Вплоть до настоящего времени практически использовались два типа АОС: конусные и рекомбинационные. Для каждого типа имеются две кинематические вариации, соответствующие соударениям в системе центра масс (обычно в контексте экспериментов по аннигиляции электронов и позитронов) и адронным соударениям. В первом случае определения подчеркивают сферическую симметрию. Во втором случае система центра масс сталкивающихся партонных неизвестна, поэтому принимаются определения в терминах поперечной энергии, являющейся инвариантом относительно бустов вдоль оси пучков. (В экспериментах по глубоконеупругому рассеянию лептонов на нуклонах имеются дополнительные особенности, см. обзор [12].) Для простоты в дальнейшем ограничимся сферически-симметричным случаем кинематики столкновений в системе центра масс. Второй случай рассматривается аналогично (см. оригинальные работы [4, 5]).

Использование именно этих двух типов АОС обусловлено их алгоритмической эффективностью. Другие типы (например, АОС в [13], который мы рассмотрим ниже) на практике не использовались.

Конусные алгоритмы введены в [10]. При этом струя определяется как все частицы, попадающие внутрь конуса углового радиуса R , а направление оси струи — с помощью какой-то итерационной процедуры, например минимизации энергии частиц, оставшихся за пределами всех струй (так называемая

недостающая энергия). Количество струй ограничивается требованием, чтобы оно было минимальным и достаточным, например, для того, чтобы недостающая энергия не превышала некоторый порог.

Уже здесь видно, сколько немотивированных элементов содержится в подобном определении (не считая параметра R): можно добиваться совпадения оси струи с суммарным 3-импульсом всех входящих в нее частиц; вместо полной недостающей энергии можно пытаться «покрыть» все событие минимальным числом конусов и отбрасывать струи с энергией ниже некоторого порога; наконец, неясно, что делать в случае перекрывающихся конусов. И даже если достичь соглашения по всем этим вопросам, все равно остается проблема выбора начального положения конусов: итерационные процедуры могут приводить к разным конечным конфигурациям в зависимости от начальной. Последний пункт означает, что определение, как было указано в [4], неполное.

Рекомбинационные алгоритмы были введены в контексте построения случайных генераторов частиц [15]. Здесь влияние идеи обращения адронизации (часто реализуемой в таких генераторах как каскад последовательных «распадов» участвующих частиц) видно особенно четко. В таких алгоритмах проверяются пары частиц в некоторой последовательности, и если пара удовлетворяет некоторому критерию, то она заменяется на одну псевдочастицу, и процесс начинается заново. Критерием окончания процесса может быть, например, отсутствие пары (псевдо)частиц, удовлетворяющих критерию для объединения. Струями объявляются псевдочастицы, энергия которых превышает, например, некоторый порог.

Как и в случае конусных алгоритмов, здесь может быть множество несущественных вариаций, не имеющих очевидного физического смысла (например, порядок обработки пар частиц). Главный произвол, однако, сосредоточен в выборе критерия рекомбинации. Исторически первым был критерий «Lucius» [15] (вместо громоздкой первоначальной формы ниже приводится упрощенный вариант, введенный в [16]; заметим, что форма записи критерия — например, нормировка левой части — может в разных работах различаться):

$$\frac{E_a E_b}{E_a + E_b} \sqrt{1 - \cos \theta_{ab}} < y_{\text{cut}}. \quad (2)$$

Здесь θ_{ab} — угол между направлениями двух частиц, а обе энергии нормированы на полную энергию события; y_{cut} — свободный параметр. Популярность рекомбинационная схема получила после работы [17], где был введен следующий простой критерий JADE:

$$E_a E_b (1 - \cos \theta_{ab}) < y_{\text{cut}}. \quad (3)$$

Левая часть имеет простой смысл инвариантной массы пары частиц. Однако оказалось, что получающийся алгоритм находит слишком много псевдоструй, которые состоят из низкоэнергетичных частиц, разделенных большими угловыми расстояниями, что плохо согласуется с физической картиной адронизации (на формальном языке это проявляется в чрезмерной чувствительности к высшим поправкам теории возмущений при теоретических расчетах). Для исправления этого недостатка был введен вариант «Geneva» [18], в котором критерий имеет вид

$$\frac{E_a E_b}{(E_a + E_b)^2} (1 - \cos \theta_{ab}) < y_{\text{cut}}. \quad (4)$$

Однако наибольшую популярность получил критерий «Durham» или k_T [19]:

$$\min(E_a, E_b) \sqrt{1 - \cos \theta_{ab}} < y_{\text{cut}}. \quad (5)$$

Аргументом в его пользу является удобство его использования для теоретических вычислений в рамках метода ведущего логарифмического приближения. Он также устраняет указанную трудность с критерием «Geneva». Наконец, в работе был выведен важный для нашего дальнейшего обсуждения «оптимальный» критерий

$$\frac{E_a E_b}{E_a + E_b} (1 - \cos \theta_{ab}) < y_{\text{cut}}, \quad (6)$$

по счастливой случайности оказавшийся средним геометрическим критериев JADE и «Geneva».

Построение псевдочастицы по паре частиц также допускает произвол: можно складывать 4-импульсы частиц, а можно 3-импульсы, но настаивать на безмассовости 4-импульсов получающихся псевдочастиц (которые в итоге должны стать струями, соответствующими партонам, а ведь партонны считаются безмассовыми). Все эти усложнения — как и вариации, отличающиеся порядком рекомбинации пар частиц, и т. п. — представляют лишь исторический интерес в свете нашей теории и здесь рассматриваться не будут.

Отметим интересную алгоритмическую вариацию, предложенную в [20]. Она заключается в том, чтобы сначала (до выполнения рекомбинаций) отметить пары частиц, удовлетворяющие рекомбинационному критерию, и затем найти связанные компоненты в графе, вершинам которого соответствуют частицы, а ребрам — отмеченные пары. Связные компоненты и объявляются струями. От такой схемы можно ожидать большей стабильности в отношении малых эффектов типа ошибок измерения по сравнению с попарными рекомбинациями, однако распространения она не получила.

Рекомбинационные алгоритмы сравнивались в работе [16] (оптимальный критерий (6) в сравнение не был включен, но это и не важно; см. ниже).

При этом использовался целый ряд разных способов сравнения и было найдено, что критерий JADE всегда оказывается наихудшим, критерий «Geneva» часто тоже плох, хотя иногда и показывает лучшие результаты, а все остальные критерии (обычно считающиеся предпочтительными) занимают полосу между этими двумя, причем разброс между «хорошими» критериями (иногда отличающимися, на первый взгляд, лишь незначительно) не так уж и мал. Отсюда легко сделать вывод для оптимального критерия: являясь средним геометрическим двух «плохих» критериев JADE и «Geneva», он будет демонстрировать поведение, «среднее» по отношению к ним и тем самым окажется среди «хороших» критериев.

Разница между различными АОС оказывается физически существенной: в работе [21] обнаружено, что основной вклад в неопределенность измерения массы топ-кварка на строящемся ускорителе LHC даст именно произвол в АОС.

Разница в алгоритмических схемах конусных и рекомбинационных алгоритмов проявляется прежде всего в том, что геометрическая форма струй, которые находят с помощью конусных алгоритмов, в основном регулярна и проста (за исключением случаев перекрывания конусов), и поэтому легче учитывать эффекты детектора (в принципе, следовало бы включить модель детектора в программу генерации событий, но практически это нереально, поэтому стремятся найти простой способ учесть подобные искажения, например, путем внесения поправок к энергиям струй и т.п.). Форма струй, которые находят с помощью рекомбинационных алгоритмов, простой быть не должна даже для случаев хорошо разделенных струй. Кроме того, результаты применения конусных алгоритмов, вследствие их геометрической жесткости, имеют тенденцию быть более устойчивыми к незначительным вариациям (например, ошибкам измерения энергий адронов), чем в случае рекомбинационных схем. Другими словами, экспериментаторы, занимающиеся обработкой данных, предпочитают использовать конусные алгоритмы, а теоретики, занимающиеся расчетами в рамках квантовой хромодинамики, — рекомбинационные (особенно k_T).

В целом можно заключить, что требование ИК-конечности само по себе оказывается недостаточным ограничением на выбор АОС, а остающийся произвол физически существен.

2. КРИТИКА «ДИНАМИЧЕСКОГО» ПОДХОДА. КИНЕМАТИЧЕСКИЕ АСПЕКТЫ

Систематический анализ трудностей построения АОС был предпринят в [2]. Логика рассуждений была такова. Прежде всего отметим, что если незначительные (в любом смысле) изменения приводят к значительным вари-

ациям результатов, то это называется неустойчивостью. Поэтому первый шаг должен состоять в том, чтобы проанализировать математическую структуру задачи на предмет корректного определения соответствующих понятий сходимости и непрерывности. Здесь нужно рассмотреть вопросы типа структуры детектора, соответствующей структуры ошибок измерения, математической природы измеряемых объектов и т. п. Необходимо также выяснить, какова предпочтительная структура наблюдаемых с точки зрения фундаментального формализма квантовой теории поля и как описываются различные физические характеристики событий в терминах таких наблюдаемых. Все это чисто кинематические вопросы (и, заметим, эксперты, которые занимались проблемой АОС, интересуясь в основном динамикой КХД, оказались не готовы к такому взгляду на проблему). В работе [2] были установлены следующие факты.

Во-первых, отображение вида (1) не может быть непрерывным, если оно не является тождественным (а оно обычно не может быть таким: число частиц в левой части порядка сотен, а число струй в правой — порядка единиц). Поэтому любая наблюдаемая, построенная как функция правой части, обязательно будет разрывной функцией на исходных адронных состояниях.

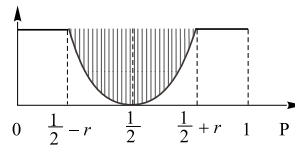
В то же время легко увидеть, что любые разрывы являются своего рода «усилителями» флуктуаций значений наблюдаемой: достаточно рассмотреть простейший пример случайной величины («состояния») P , равномерно распределенной на единичном интервале $(0,1)$, и вычислить вариации значений разрывной функций $f_{\text{разр}}(P) = \theta(P - 0,5)$ и любого ее непрерывного аналога, отличающегося тем, что разрыв тем или иным образом «регуляризован», скажем, посредством линейной интерполяции на интервале, симметричном относительно разрыва:

$$f_{\text{пер}}(P) = \begin{cases} 0, & \text{если } P < \frac{1}{2} - r, \\ 1, & \text{если } P > \frac{1}{2} + r, \\ \frac{P - 1/2 + r}{2r}, & \text{если } \frac{1}{2} - r < P < \frac{1}{2} + r. \end{cases} \quad (7)$$

Это видно, если посмотреть на поведение выражения $(f(P) - \langle f \rangle)^2$, интеграл от которого дает вариацию (см. рисунок).

Увеличение вариации наблюдаемой за счет разрывов (соответствует заштрихованной площади на рисунке) приводит к увеличению ошибок измерения физических параметров.

Важно заметить, что любое двузначное (да/нет) решение о рекомбинации пары (псевдо)частиц в одну в соответствии с критериями



типа (2)–(6) приводит к новому разрыву, расположенному на некотором достаточно сложном многообразии в пространстве исходных состояний. Итерационный характер рекомбинаций может приводить к тому, что искажения конфигурации (псевдо)частиц, возникающие на каждой итерации, накапливаются. В этом смысле следует ожидать, что рекомбинационные алгоритмы являются более чувствительными, чем, скажем, конусные алгоритмы, к малым вариациям типа ошибок измерения энергий частиц.

С другой стороны, регуляризация разрывных наблюдаемых оказывается новым, до сих пор не использовавшимся путем повышения качества процедур обработки данных, причем весьма простым (и дешевым в сравнении со стоимостью усовершенствований экспериментальной аппаратуры, способных дать тот же эффект).

Во-вторых, точка зрения на АОС как на обращение адронизации может быть адекватной, только если пренебрегать поправками теории возмущений. В самом деле, как невозможно сказать, через какое отверстие прошел электрон в знаменитом мысленном квантово-механическом эксперименте, точно так же нельзя с полной уверенностью сказать, например, из какого партона в жестком состоянии «излучилась» та или иная наблюдаемая частица, особенно, если соответствующие струи близки, а частица находится между ними. Другими словами, обращение адронизации — строго говоря, некорректная операция, и это особенно важно в ситуациях, когда необходимо учитывать радиационные поправки. Поэтому для систематического анализа АОС, т. е. такого анализа, который не опирается на интуицию, связанную с нижшим порядком теории возмущений, требуется найти другое основание.

В-третьих, теоретические расчеты гораздо дешевле экспериментальных данных и в гораздо меньшей степени ограничены материальными обстоятельствами. Поэтому рациональная постановка вопроса должна состоять в том, чтобы искать прежде всего такой АОС, который позволял бы максимизировать физическую информацию, извлекаемую из экспериментальных данных. При этом удобство теоретиков — вопрос вторичный.

Максимизировать извлекаемую информацию — значит, минимизировать ошибки измеряемых параметров (таких как масса топ-кварка). Это дает начальную точку для систематического вывода оптимального АОС. Однако чтобы такой вывод стал возможным, нужно формализовать ряд важных понятий.

3. ТЕОРИЯ *C*-НЕПРЕРЫВНЫХ НАБЛЮДАЕМЫХ

Обратимся к конструктивным аргументам [2] и отошлем за обоснованиями к оригинальному тексту.

Во-первых, в работе [2] был выяснен ряд вопросов, освещающих математическую природу задачи определения адронных струй в свете проблемы построения оптимальных наблюдаемых для измерения физических параметров.

Событие, наблюдаемое в мультимодульном калориметрическом детекторе, с математической точки зрения следует интерпретировать как положительную меру на единичной сфере. Другими словами, интерпретировать событие в левой части выражения (1) следует как

$$P(p) = \sum_{i=1}^{N_{\text{part}}} E_i \delta_{\hat{\mathbf{p}}_i}(\hat{p}), \quad (8)$$

где E_i — энергия i -й частицы; \hat{p} — точка на единичной сфере; а $\hat{\mathbf{p}}_i$ — единичный вектор направления i -й частицы. При этом δ -функция определена равенством

$$\int d\hat{p} f(\hat{p}) \delta_{\hat{\mathbf{p}}_i}(\hat{p}) = f(\hat{\mathbf{p}}_i) \quad \text{для любой непрерывной } f(\hat{p}). \quad (9)$$

Здесь интегрирование идет по единичной сфере, $d\hat{p}$ — инфинитезимальный элемент площади поверхности.

Более того, структура калориметрических детекторов однозначно диктует определенный тип сходимости для событий, представленных в виде (8), — так называемую слабую сходимость. Точнее говоря, последовательность событий P_n сходится к некоторому событию P , если для каждой непрерывной функции $d(\hat{p})$ имеет место обычная сходимость соответствующей числовой последовательности

$$\int d\hat{p} d(\hat{p}) P_n(\hat{p}) \xrightarrow{n \rightarrow \infty} \int d\hat{p} d(\hat{p}) P(\hat{p}). \quad (10)$$

Фактически это происходит потому, что интегралы типа (10) моделируют элементарные калориметрические модули (при этом d принимает значения между 0 и 1). Непрерывный характер изменения значений d соответствует неабсолютной эффективности регистрации для частиц, попадающих в граничные области модуля. Формула (10) означает, что события становятся неразличимыми, если их не может различить никакая калориметрическая ячейка.

Сходимость (10) была названа в [2] C -сходимостью (calorimetric). Ее можно выразить на языке некоторой функции расстояния (см. разд. 3 в [4]), однако такое выражение имеет искусственный характер.

Вообще говоря, на бесконечномерных пространствах (каким является и пространство событий (8) — количество частиц в пределе высоких энергий ограничить нельзя) возможны разные сходимости. При этом принятие разных

типов сходимости индуцирует разные математико-алгоритмические конструкции. Например, в ряде ранних работ [22] пространство событий наделялось квадратичной метрикой, превращавшей его в гильбертово пространство, а это привело к рассмотрению определенного типа наблюдаемых (связанных с системой ортогональных функций на сфере) и в конечном счете направило анализ по ложному пути.

Следует четко понимать, что структура ошибок каждого типа измерительных приборов (в нашем случае — многомодульных калориметрических детекторов) однозначно фиксирует структуру измеряемых объектов (в нашем случае — многочастичных событий) *вместе* с некоторой сходимостью в пространстве этих объектов. Это подобно тому, как измеряемая длина представляется рациональным числом, причем сходимость в «пространстве длин» — это обычная сходимость вещественных чисел, а не, скажем, сходимость p -адических чисел. Принять первую нас заставляет именно структура ошибок измерения соответствующих измерительных приборов (линеек). Ясно, насколько бессмысленными были бы в механике приближенные методы, построенные на основе p -адической сходимости чисел.

Во-вторых, из приведенных рассуждений о преимуществе непрерывных наблюдаемых немедленно следует простой эвристический «рецепт» повышения качества обработки данных. Действительно, разрывы наблюдаемых можно рассматривать как «щели», через которые «теряется» физическая информация (на самом деле разрывы скорее служат «усилителями» статистических флуктуаций в исходных данных, но конечный эффект тот же: уменьшение отношения сигнал/шум). Поэтому простейший способ действий состоит в устранении (регуляризации) таких разрывов: любое «увеличение непрерывности» устраняет эффект усиления флуктуаций вблизи разрыва («вырезание» заштрихованной области на рисунке). Во многих случаях такие регуляризации очень легко строить. Действительно, экспериментальные процедуры, как правило, состоят из многоэтапного отбора событий, причем каждый этап регулируется критерием вида

$$\phi(P) > c_{\text{cut}}, \quad (11)$$

где ϕ — некоторая более-менее непрерывная функция события; c — параметр, фиксирующий обрезание. Такой отбор эквивалентен тому, как если бы конечная наблюдаемая считалась определенной на любых событиях, но в нее вводился бы фактор

$$\Theta(P) = \theta(\phi(P) - c_{\text{cut}}), \quad (12)$$

где θ — обычная функция Хевисайда, принимающая значения 0 и 1. Тогда регуляризация заключалась бы в замене разрывной θ -функции на какой-то ее

непрерывный аналог, например,

$$\Theta_{\text{reg}}(P) = \begin{cases} 1, & \text{если } \phi(P) > c_+, \\ \frac{\phi(P) - c_-}{c_+ - c_-}, & \text{если } c_- < \phi(P) < c_+, \\ 0, & \text{если } \phi(P) < c_-. \end{cases} \quad (13)$$

Указан лишь простейший линейный тип интерполяции между двумя крайними значениями (0 и 1) в некотором интервале $c_- < c_{\text{cut}} < c_+$. Возможен и другой путь:

$$\Theta_{\text{reg}}(P) = \int_{c_-}^{c_+} dc w(c) \theta(\phi(P) - c), \quad (14)$$

где w — некоторый вес. При $w = 1$ получается формула (13). Именно эта конструкция фигурирует в наблюдаемой (17), обсуждаемой ниже.

Таким образом, событиям приписывается нецелый вес, и в окончательном выражении для наблюдаемой следует учитывать этот вес при суммировании вкладов событий. Другими словами, если представить наблюдаемую, соответствующую некоторой процедуре обычного типа в виде $\sum_i f(P_i)$, где суммирование идет только по «отобранным» событиям, т.е. удовлетворяющим условию (11), то регуляризация вызывает модификацию следующего вида:

$$\sum_i f(P_i) \equiv \sum_j \theta(\phi(P) - c_{\text{cut}}) f(P_j) \xrightarrow{\text{регуляризация}} \sum_j \Theta_{\text{reg}}(P) f(P_j), \quad (15)$$

где в средней и правой суммах учитываются все события с $\Theta_{\text{reg}}(P) > 0$ (которых, вообще говоря, больше, чем в самой левой сумме).

Этот эвристический «рецепт» является весьма общим и, по крайней мере в некоторых случаях, может дать значительный эффект (см. ниже обсуждение наблюдаемой (17)). При этом особенно ценно то, что стоимость регуляризации пренебрежимо мала по сравнению с эквивалентным (по достигаемому увеличению отношения сигнал/шум) усовершенствованием аппаратуры.

В-третьих, понятно, что если требовать непрерывности наблюдаемых, оптимизирующих ошибки измерения физических параметров, то, строго говоря, наблюдаемые должны быть непрерывными функциями состояний именно относительно C -сходимости. Сразу возникают два вопроса: 1) какова структура таких наблюдаемых? 2) можно ли — и если можно, то как — построить C -непрерывные наблюдаемые для измерения параметров типа массы нестабильной частицы, продукты распада которой содержат струи? Конструктивные ответы на эти вопросы были даны в работе [2] и уточнены в [4].

Простейшими C -непрерывными наблюдаемыми являются выражения типа правой части (10) с непрерывными d , что следует из определения.

Чуть более сложными являются хорошо известные наблюдаемые, называемые «shape observables», такие как «thrust», «sphericity», «acomplanarity» и т. п. [23]. Применение таких наблюдаемых ограничено: их в основном используют при измерении константы связи КХД α_S (см., например, [24]) либо как вспомогательные величины (впрочем, не слишком полезные) для нахождения критериев отбора событий с помощью, например, нейронных сетей (см. [25]). Все эти наблюдаемые являются линейными по энергиям частиц.

Более сложные наблюдаемые, полилинейные по энергиям частиц, имеют вид

$$\sum_{i_1} \dots \sum_{i_N} E_{i_1} \dots E_{i_N} f(\hat{\mathbf{p}}_{i_1}, \dots, \hat{\mathbf{p}}_{i_N}), \quad (16)$$

где f — симметричная и непрерывная функция своих аргументов (N единичных векторов), а суммирования идут по всем частицам состояния. Интересно, что в работе [22] рассматривались бесконечные серии переменных такого типа, но лишь специального вида, мотивированного произвольным введением гильбертовой нормы в пространстве наблюдаемых. Извлечь существенную пользу из конструкций такого рода в [22] не удалось.

Далее, в [2] был указан ряд конструкций с наблюдаемыми вида (16), с помощью которых можно получить большое разнообразие функций, чувствительных к физически важным характеристикам состояний типа инвариантных масс двухструйных подсистем, причем для этого не нужно применять никакого АОС, и ни на каком шаге построения не происходит нарушения свойства C -непрерывности. Набор таких конструкций был назван C -алгеброй. Нетривиальный момент здесь состоит в том, что на языке C -непрерывных наблюдаемых удалось выразить практически все свойства состояний, обычно описываемые посредством АОС. В работе [2] подчеркивалось, что широко используемая классификация состояний по числу струй является важнейшим источником неоптимальности обычных процедур обработки данных в отношении точности измеряемых физических параметров. Это имеет место потому, что число струй — разрывная целочисленная наблюдаемая, причем разрывы находятся в областях быстрого изменения вероятностей и поэтому возникают нестабильности относительно радиационных поправок, связанных с излучением партонов под небольшими углами. Затем в [2] были даны примеры наблюдаемых, несущих ту же физическую информацию, что и, скажем, сечение рождения трехструйных событий, но обладающих свойством C -непрерывности. Такие наблюдаемые с необходимостью принимают нецелые значения и интерпретируются как «нецелое число струй».

Практическая полезность всех этих идей была продемонстрирована в работе [25], где дается обзор экспериментальных процедур, использовавшихся для открытия и изучения свойств топ-кварка в FNAL (Батавия, США). Как

известно, самым трудным для наблюдения топ-кварка является чисто адронный канал распада $t\bar{t} \rightarrow 6$ струй из-за большого фона от чисто адронных процессов и большого числа струй, так что вероятность конфигураций с перекрывающимися струями повышается. В коллаборации D0 существенного прогресса с выделением сигнала в этом канале удалось добиться введением непрерывной переменной (хотя и построенной на основе некоторого АОС), интерпретируемой как нецелое число струй:

$$N_{\text{jets}}^A = \int_{15 \text{ ГэВ}}^{55 \text{ ГэВ}} E_T N(> E_T) dE_T / \int_{15 \text{ ГэВ}}^{55 \text{ ГэВ}} E_T dE_T, \quad (17)$$

где $N(> E_T)$ — число струй с энергией больше E_T (речь идет об адронных столкновениях и все энергии берутся поперечными, т. е. инвариантными относительно бустов вдоль оси пучков). Переменная (17) представляет собой регуляризованную версию целочисленной переменной «число струй с поперечной энергией больше некоторого $E_{T,\text{cut}}$ » (ср. (14)). В данном случае выбор веса (E_T) подчеркивает вклады более энергичных струй. Для целей регуляризации было бы достаточно использовать, скажем, единичный вес вместо указанного; из работы [26] неясно, из каких соображений был сделан конкретный выбор. Итак, среди всех наблюдаемых, которые возникают при изучении многоадронных состояний, выделяются следующие три класса.

1. Элементарные наблюдаемые, описываемые выражениями типа правой части (10) (для реальных событий они сводятся к однократным суммам вида (16) с $N = 1$). Это фундаментальные наблюдаемые, соответствующие простейшим калориметрическим ячейкам.

2. Корреляторы вида (16). Такие наблюдаемые также играют центральную роль в теории ввиду их связи с формализмом квантовой теории поля (см. разд. 4).

3. Общие C -непрерывные наблюдаемые. Способы построения таких наблюдаемых для описания различных свойств адронных состояний, обычно выражаемых на языке струй, даны в [2]. Как отмечалось в [4], имеет место математический факт, состоящий в том, что произвольные C -непрерывные наблюдаемые получаются предельным переходом из конечных алгебраических комбинаций элементарных наблюдаемых или, что эквивалентно, из корреляторов (16). Это следует из известной теоремы Банаха–Алаоглу (см., например, [27]).

Уже здесь возникает идея о том, что АОС можно трактовать как способ упрощенного построения C -непрерывных наблюдаемых, поскольку их точное вычисление на событиях с сотней и более частиц может быть чрезвычайно дорогостоящим. Другими словами, подходящие C -непрерывные наблюдаемые строятся как функции конфигураций струй, соответствующих событиям. Тогда ясно, что конфигурацию струй-псевдочастиц (правая часть (1)) следует

выбирать так, чтобы она была максимально близка к исходной многочастичной конфигурации (левая часть (1)) *в смысле C-сходимости*. Поскольку такая сходимость определяется в терминах близости значений простейшего класса функций, определенных на конфигурациях (псевдо)частиц (см. (10)), то достаточно, чтобы АОС (1) минимизировал разность

$$\left| \int d\hat{p} d(\hat{p}) Q(\hat{p}) - \int d\hat{p} d(\hat{p}) P(\hat{p}) \right| = \left| \sum_a E_a d(\hat{p}_a) - \sum_j E_j d(\hat{p}_j) \right|. \quad (18)$$

Эта задача уже поддается аналитическому изучению.

Подчеркнем, что в такой трактовке АОС выглядит лишь как вычислительный прием, что может, на первый взгляд, показаться не соответствующим физическому смыслу, обычно связываемому со взглядом на струи как наблюдаемые «следы» кварков и глюонов. Однако никакого противоречия здесь нет: «физический смысл» представляет собой описание качественного поведения соответствующих вероятностей (на специфическом языке), а успех любой аппроксимации непосредственно определяется тем, насколько хорошо она отражает такое качественное поведение. Поэтому, требуя адекватности АОС как вычислительного приема, мы, пусть и косвенно, требуем, чтобы в нем отражались наиболее существенные динамические черты изучаемых физических процессов.

Однако прежде чем переходить к построению оптимального АОС, отвечающего требованию минимизации разности (18), полезно рассмотреть еще несколько вопросов, углубляющих теорию, которая приводит к этому требованию.

4. КВАНТОВО-ПОЛЕВЫЕ АСПЕКТЫ. КВАРК-АДРОННАЯ ДУАЛЬНОСТЬ

Рассмотрим теперь следующий вопрос. Если речь идет о высокоточных теоретических вычислениях, то такие вычисления легче делать, если структура наблюдаемых соответствует структуре теоретического формализма. В данном случае таким формализмом является квантовая теория поля, в рамках которой системы с переменным (причем обычно большим и неопределенным) числом частиц описываются с помощью разного рода корреляторов. Поэтому естественно стремиться выразить физику многоадронных состояний на языке корреляторов вида (16). В работе [3] был установлен следующий фундаментальный факт: корреляторы вида (16) допускают точное выражение через корреляторы плотности тензора энергии-импульса — фундаментального объекта в формализме квантовой теории поля.

Перепишем среднее значение наблюдаемой (16) в следующем виде:

$$\begin{aligned} & \langle \sum_{i_1} \dots \sum_{i_N} E_{i_1} \dots E_{i_N} f(\hat{\mathbf{p}}_{i_1} \dots \hat{\mathbf{p}}_{i_N}) \rangle_{\mathbf{P}} = \\ & = \int d\mathbf{n}_1 \dots \int d\mathbf{n}_N \langle in | \epsilon(\mathbf{n}_1) \dots \epsilon(\mathbf{n}_N) | in \rangle \times f(\mathbf{n}_1, \dots, \mathbf{n}_N), \end{aligned} \quad (19)$$

где \mathbf{n}_i — единичные векторы, а операторная плотность $\epsilon(\mathbf{n})$ на единичной сфере задается выражением

$$\epsilon(\mathbf{n}) = \int \frac{d\mathbf{p}}{2p^0} a^+(\mathbf{p}) a^-(\mathbf{p}) \times (\mathbf{p}\mathbf{n}) \delta(\hat{\mathbf{p}}\mathbf{n}). \quad (20)$$

Тогда имеет место следующий факт [3]:

$$\epsilon(\mathbf{n}) d\mathbf{n} = \lim_{t \rightarrow +\infty} \int_{\mathbf{x} \in \text{Cone}(t, \mathbf{n}, d\mathbf{n})} dx n_i T_{0i}(\mathbf{x}, t) d\mathbf{n}, \quad (21)$$

где $\text{Cone}(t, \mathbf{n}, d\mathbf{n})$ — конус радиуса t с инфинитезимальным угловым раствором $d\mathbf{n}$ около оси \mathbf{n} , а $T_{ij}(\mathbf{x}, t)$ — пространственно-временная плотность оператора тензора энергии импульса.

Значение результата (21) состоит в том, что он подводит более твердое основание под гипотезу кварк-адронной дуальности, т. е. возможности теоретических предсказаний для адронов, полученных в рамках теории возмущений для кварков и глюонов. В самом деле, хотя исходное выражение сформулировано в терминах наблюдаемых (т. е. адронных) асимптотических полей, явно фигурирующих в уравнении (20), правая часть уравнения (19) вместе с выражением (21) не содержит никакого намека на структуру асимптотических полей. Поэтому представляется естественным и достаточным (по крайней мере для неадронных состояний in) апеллировать к свойству асимптотической свободы КХД, чтобы с уверенностью развивать теорию возмущений для корреляторов вида (19).

Заметим, что ввиду установленной структуры семейства C -непрерывных наблюдаемых требование ИК-конечности для них сводится (в предположении гладкости угловых функций f) к существованию в рамках теории возмущений плотности тензора энергии-импульса как хорошо определенной величины.

Важное следствие состоит в возможности систематическим образом исследовать структуру степенных поправок к теоретическим предсказаниям для рассматриваемых корреляторных наблюдаемых в случае неадронных состояний in (случай электрон-позитронной аннигиляции в адроны). Для этого достаточно следовать общему подходу, восходящему к методу правил сумм КХД [28] и систематически сформулированному в [29]. Применение подхода

[29] основано на возможности получать теоремы факторизации (обобщения операторного разложения) в пространстве Минковского, открытой при распространении метода асимптотической операции (см. обзор [30]) на произвольные асимптотические режимы [31].

5. СТАТИСТИЧЕСКИЕ АСПЕКТЫ

Итак, установлено, что использование определенных типов (а именно C -непрерывных) наблюдаемых предпочтительно с точки зрения максимального извлечения физической информации из наблюдаемых событий. Однако это понимание позволяет только улучшать уже заданные наблюдаемые посредством регуляризации разрывов. Возникает естественный вопрос: существует ли регулярный — пусть и приближенный — способ построения наблюдаемых, по возможности оптимальных в плане точности извлечения физической информации из данных? Положительный и конструктивный ответ на этот вопрос найден в работах [4, 6] в форме так называемого метода квазиоптимальных наблюдаемых, который и описывается в этом разделе.

С точки зрения математической статистики в задачах рассматриваемого типа в конечном счете речь идет об оценивании параметров некоторого статистического распределения (на пространстве многоадронных событий). Извлечение параметра путем фитирования значения некоторой наблюдаемой — это не что иное, как обобщенный метод моментов, описанный в любом учебнике математической статистики для физиков (см., например, [32]). Удивительно то, что имеющиеся учебники не пытаются дать ответ на следующий простой вопрос: при данной плотности вероятности $\pi(P)$, зависящей от параметра M , какой обобщенный момент (наблюдаемая) $f(P)$ приводит к минимальной вариации для соответствующей оценки M по методу обобщенных моментов? Заметим, что знание такого ответа дает простой и общий рецепт построения «хороших» наблюдаемых, допускающий множество алгоритмических реализаций (мы будем называть этот рецепт методом квазиоптимальных наблюдаемых [4, 6]). При том что в больших экспериментах используются практически все известные статистические методы, включая весьма изощренные типа нейронных сетей (см., например, обзор [25]), отсутствуют какие-либо указания на то, что простой, мощный и полезный метод квазиоптимальных наблюдаемых известен специалистам по обработке данных. Опишем его, отсылая за обоснованиями к оригинальным работам [6, 4].

Любая точечная оценка θ параметра M по случайной выборке из N событий имеет некоторую дисперсию $N^{-1} \text{Var } M(\theta)$ (при больших N), про которую известно, что она не может быть меньше некоторой величины, не зависящей от метода оценки (неравенство Крамера–Рао; см. [32]). Поэтому в случае оценок по методу обобщенных моментов естественно попытаться

определить точку минимума вариации как функцию обобщенного момента (наблюдаемой) f . В этом случае можно записать явное выражение

$$\text{Var } M(f) = \left(\frac{\partial \langle f \rangle}{\partial M} \right)^{-2} \text{Var } f, \quad \text{где } \text{Var } f = \langle f^2 \rangle - \langle f \rangle^2 \quad (22)$$

и затем попытаться найти точку и значение его минимума. Простые вычисления (функциональный аналог разложения Тейлора), справедливые практически при тех же условиях, что и метод максимального правдоподобия, показывают, что минимум достигается на наблюдаемой вида

$$f_{\text{opt}}(P) = \frac{\partial \ln \pi(P)}{\partial M} \quad (23)$$

с точностью до произвольных аддитивной и ненулевой мультипликативной констант. При этом около точки минимума

$$\text{Var } M(f_{\text{opt}} + \phi) = \frac{1}{\langle f_{\text{opt}}^2 \rangle} + \frac{1}{\langle f_{\text{opt}}^2 \rangle^3} \{ \langle f_{\text{opt}}^2 \rangle \times \langle \bar{\phi}^2 \rangle - \langle f_{\text{opt}} \times \bar{\phi} \rangle^2 \} + \dots, \quad (24)$$

где $\bar{\phi} = \phi - \langle \phi \rangle$, а значение в минимуме в точности соответствует неравенству Крамера–Рао. Положительность выражения в фигурных скобках следует из неравенства Буяковского–Коши–Шварца. Похожие квадратичные оценки встречаются в стандартных доказательствах неравенства Крамера–Рао, однако их интерпретации внимания не уделяется.

Формула (24) красноречива: она показывает, что отклонения вариации от минимального значения квадратичны по отклонению наблюдаемой от оптимальной (величина $\bar{\phi} = \phi - \langle \phi \rangle$). Это означает, что достаточно построить приближенное выражение для оптимальной наблюдаемой f_{opt} — квазиоптимальную наблюдаемую f_{quasi} . Соответствующая малость контролируется выражением в фигурных скобках в правой части (24).

В задачах физики высоких энергий, о которых идет речь в данном обзоре, теоретическое выражение для плотности вероятности в аналитическом виде неизвестно (так что метод максимального правдоподобия применить трудно), но имеется соответствующий генератор случайных событий, и этого в принципе достаточно, чтобы строить квазиоптимальные наблюдаемые приближенно по формуле (23). За исключением дифференцирования по M , задача здесь подобна тому, что делается в алгоритмах адаптивного интегрирования Монте-Карло, которых имеется большое разнообразие. Более того, явный аналитический вид и простой смысл формул (23), (24) позволяет использовать различные эвристические подходы.

Простейший эвристический подход (непосредственно имеющий отношение к задаче об АОС) заключается в том, чтобы начинать с кусочно-постоянной f_{quasi} (в большинстве реальных ситуаций f_{opt} будет иметь не

слишком сложное поведение). Такая аппроксимация могла бы затем служить искомым начальным пунктом для применения регуляризующих процедур, обсуждавшихся выше. В контексте адронных струй такая кусочно-постоянная аппроксимация заключалась бы в том, что f_{quasi} выбирается совпадающей на всех событиях, дающих одинаковый набор струй при некотором выборе АОС. Другими словами, f_{quasi} сначала ищется в виде функции параметров струй, а не всего набора параметров исходного события. Здесь сразу возникает вопрос о том, какой АОС наиболее подходит для этой цели. Можно показать [4], что АОС должен быть таким, чтобы конфигурация струй (правая часть (1)) была бы максимально близка к исходной конфигурации частиц (левая часть (1)) в смысле C -сходимости (10), и тогда мы возвращаемся к требованию, чтобы АОС минимизировал разности (18).

6. ОПТИМАЛЬНОЕ ОПРЕДЕЛЕНИЕ СТРУЙ

Видно, что различные пути рассуждений приводят к требованию, чтобы оптимальный АОС минимизировал разности (18) для достаточно регулярных функций d , определенных на единичной сфере (ниже мы увидим, как решается проблема одновременной оценки выражения (18) для разных d). Вывести отсюда аналитическую форму оптимального АОС можно, грубо говоря, выполнив разложение в ряд Тейлора по угловым расстояниям между направлениями частиц \hat{p}_a и струй \hat{q}_j . Технические трудности здесь заключаются в том, что, во-первых, неизвестно, какой струе должна принадлежать та или иная частица, и, во-вторых, отдельно должны трактоваться частицы с низкой энергией, находящиеся на больших угловых расстояниях от всех струй, при том что положение струй не предполагается заранее известным.

Прием, позволяющий обойти эти трудности, заключается во введении новых переменных — рекомбинационной матрицы z_{aj} , эвристически интерпретируемой как доля энергии частицы с номером a , которая идет на формирование струи номер j (эта интерпретация не является изначально необходимой, но вполне естественно возникает «a posteriori»). В процессе вывода именно матричные элементы z_{aj} займут место истинных неизвестных величин, которые ищет алгоритм минимизации. Удобно заранее наложить ограничения

$$z_{aj} \geq 0, \quad \sum_j z_{aj} \equiv 1 - \bar{z}_a \leq 1. \quad (25)$$

Во втором ограничении следует предпочесть неравенство равенству, поскольку, как уже говорилось, нужно разрешить некоторым частицам вообще не участвовать в образовании струй. Подобная опция имеется во всех обычных АОС (слишком мягкие струи отвергаются, поскольку они по физическим соображениям не могут быть связаны с жесткими партонами). В рамках описы-

ваемой теории эта опция совершенно естественна, т. к. отбрасывание частиц с достаточно малыми энергиями является столь же законным способом приближенного вычисления C -непрерывных функций, как и группировка частиц в струи (ряд примеров дан в [2]).

Вывод начинается со следующего простого неравенства для разности (18), в котором использована матрица (25):

$$\left| \sum_a E_a d(\hat{p}_a) - \sum_j E_j d(\hat{p}_j) \right| \leq \sum_a \bar{z}_a E_a |d(\hat{p}_a)| + \sum_j \left| \sum_a z_{aj} E_a d(\hat{p}_a) - E_j d(\hat{q}_j) \right|. \quad (26)$$

Видно, почему рассуждения ведутся о разбиении частиц на фрагменты, относящиеся к струям, а не наоборот: симметрия нарушена тем, что число струй предполагается много меньше, чем число частиц, так что желательно учесть сокращения между как можно большим числом вкладов (крайне правая сумма по a). Далее в каждой разности в правой части производится разложение Тейлора функции d по \hat{p}_a около \hat{q}_j . Из требований сокращения ведущих и первых степеней угловой переменной довольно неожиданно возникают соотношения, эквивалентные законам сохранения энергии и 3-импульса, которые удобно записать в релятивистски-ковариантном виде:

$$q_j = \sum_a z_{aj} p_a. \quad (27)$$

Таким образом, рекомбинационная матрица осталась единственным неизвестным, определяющим все параметры струй.

Остаток без потери общности можно ограничить сверху (см. детали вклада в [4]) выражением $c_d \Omega_R$, где

$$\Omega_R = R^{-2} Y + E_{\text{soft}}, \quad (28)$$

причем Y и E_{soft} зависят не от d , а только от события, и имеют следующий вид:

$$E_{\text{soft}}(P, Q) = \sum_a \bar{z}_a E_a, \quad Y(P, Q) = \sum_j Y_j(P, Q) \equiv 2 \sum_j q_j \tilde{q}_j, \quad (29)$$

где \tilde{q}_j — светоподобный 4-импульс с единичной энергией, 3-импульс которого в сферически-симметричном случае направлен по пространственной компоненте 4-вектора q_j .

Не зависящий от события и z_{aj} коэффициент перед Y в (28) остается произвольным; его обозначение связано с его ролью в окончательном алгоритме (см. ниже).

Таким образом, оказывается, что «ошибка аппроксимации» (18) контролируется довольно простым выражением (28) с единственным свободным параметром R , причем оценка сохраняется и для произвольных, достаточно регулярных C -непрерывных наблюдаемых, что позволяет интерпретировать величину (28) как контролирующую потерю физической информации при переходе от полного описания события к его описанию в терминах струй.

Поиск оптимальной конфигурации струй, минимизирующей потерю физической информации, сводится теперь к поиску рекомбинационной матрицы z_{aj} , минимизирующей выражение (28), чтобы значение минимума не превышало некоторого порога при минимальном количестве струй. То есть искомое оптимальное определение струй запишется в виде

$$\min \Omega_R \leq y_{\text{cut}}. \quad (30)$$

Положительный параметр в правой части аналогичен параметру «углового разрешения», фигурирующему в рекомбинационных алгоритмах (ср. выражения (2)–(6)). Однако преимуществом алгоритмической схемы, основанной на (30) и (28), является отсутствие потенциально дестабилизирующих повторных рекомбинаций.

Если записать критерий (30) для пары частиц, то обнаружится, что, по существу, восстанавливается «оптимальный» критерий (6). Это позволяет утверждать без детального тестирования, что полученный критерий должен давать результаты, в общих чертах подобные другим «хорошим» рекомбинационным алгоритмам (см. обсуждение после уравнения (6)).

Отметим важную роль параметра R в выражении (28): он эффективно ограничивает сверху угловой размер струй (см. подробнее о механизме этого явления в [4]). В этом отношении найденный критерий напоминает конусные алгоритмы, хотя геометрия струй в нашем критерии учитывает динамику в том смысле, что реальный радиус струи зависит от распределения энергии внутри нее (при заданном R максимальный размер струи, состоящей из двух частиц, минимален, когда энергии частиц равны, и максимален, когда почти вся энергия уносится одной из частиц). Хотя такой «динамичный» характер формы струй затрудняет непосредственное использование методик, применяемых для учета детекторных эффектов на уровне струй и разработанных для случая фиксированной конусной геометрии, все же получающиеся струи будут иметь явно более регулярную форму (выпуклый конус), чем в случае рекомбинационных алгоритмов. Кроме того, в отличие от конусных алгоритмов проблема перекрывающихся струй здесь не возникает в принципе: частицы, находящиеся в промежутке между близкими струями, относятся к одной из

струй (или исключаются из всех струй) в зависимости от того, какой из Y_j (или E_{soft}) вырастет в меньшей степени.

Принципиально важно то, что в случае определения (30) решается проблема выбора начальных условий (выбор начальной конфигурации струй в конусных алгоритмах или порядок обработки частиц в рекомбинационных): по физическому смыслу левой части (оценка потери физической информации) следует выбирать конфигурацию струй (рекомбинационную матрицу z_{aj}), для которой достигается глобальный минимум левой части. (Множество состояний с вырожденным глобальным минимумом имеет меру нуль, т.е. такие состояния реализуются с нулевой вероятностью. Здесь возможно возникновение неустойчивостей относительно малых эффектов, но здесь применима общая схема регуляризации; детали см. в [4].)

С теоретической точки зрения важно то, что левая часть в уравнении (30) представляет собой обобщение хорошо изученной наблюдаемой «thrust» (см. [11, 24]) при двух противоположно направленных полуосях (положение которых также находится с помощью некоторой оптимизации) на случай любого числа произвольно направленных полуосей (совпадающих теперь с осями струй). Такие наблюдаемые лучше всего подходят для учета высших поправок в теоретических расчетах (см. результаты, цитированные в [24]). Известно, что высшие поправки (включая даже двухпетлевые поправки) важны, как и поправки, получаемые с помощью суммирования ведущих логарифмов [24]. Поэтому с точки зрения качества теоретических вычислений связь нового критерия (30) и хорошо изученной наблюдаемой «thrust» чрезвычайно привлекательна.

Отметим, что теоретические достоинства наблюдаемых типа «thrust» были оценены еще на заре физики струй [23], и попытки использовать обобщения таких наблюдаемых для определения струй известны: в работе [13] рассматривался критерий вида $\min Y \leq y_{\text{cut}}$, получающийся из оптимального отбрасыванием компоненты E_{soft} (совпадение является случайным, т.к. критерий (30) был выведен из первых принципов полностью аналитически, «вслепую», без привлечения априорных геометрических соображений). Кроме того, там был найден лишь приближенный эвристический алгоритм поиска минимума, так что реально АОС из работы [13] свелся к постулированию некоторого специфического алгоритма. Поэтому неясно, как теоретические расчеты для идеального критерия $\min Y \leq y_{\text{cut}}$ должны соотноситься с принятым в [13] эвристическим алгоритмом, особенно в контексте прецизионных расчетов и измерений. Кроме того, наличие вклада E_{soft} в нашем случае играет важную роль, ограничивая сверху максимальный угловой радиус струй. Наконец, в отсутствие достаточно глубокого обоснования и достаточно эффективной реализации критерий, постулированный в [13], оставался столь же произвольным и уязвимым для критики, как и рекомбинационные схемы.

7. ЗАМЕЧАНИЯ ОБ АЛГОРИТМИЧЕСКОЙ РЕАЛИЗАЦИИ

При всех достоинствах найденного определения струй (30) оно было бы бесполезно без эффективного алгоритма, реализующего поиск минимума. Задача здесь сводится к поиску точки глобального минимума функции Ω_R , аргументами которой являются рекомбинационные матрицы z_{aj} , удовлетворяющие ограничениям (25). Количество матричных элементов равно числу частиц, умноженному на число струй. Например, для строящегося ускорителя ЛНС следует ожидать, что число частиц будет порядка сотни (средняя множественность адронов равна примерно 140, причем следует учитывать и возможные события с 400 частицами). Число же струй, например, в чисто адронном распаде пары кварков $t\bar{t}$, равно 6 (без учета излучения дополнительных глюонов). Таким образом, речь идет о задаче поиска глобального минимума в области размерности $O(1000)$ с примерно таким же числом ограничений. Это обстоятельство само по себе подкрепляет пессимизм, выраженный в работе [16], посвященной сравнению различных АОС, по поводу возможности организовать эффективный перебор частиц на предмет их группировки в струи в алгоритме минимизации работы [13], аналитическая форма которого близка к оптимальному определению (30). Однако оказалось, что такой эффективный алгоритм существует [5]. Кратко опишем его.

О минимизируемой функции Ω_R имеется много аналитической информации: она гладкая, довольно простая, сложный рельеф (узкие овраги и т. п.) отсутствует, выражения для производных могут быть найдены аналитически и имеют простой вид. Однако численные эксперименты показали, что минимум всегда достигается на границе области (в точке минимума z_{aj} всегда принимают значения 0 или 1); более того, любой вариант поиска, основанный на использовании градиентов, быстро выводит на границу области. В результате метод наискорейшего спуска [33] оказался непригоден. Фактически траектория поиска минимума в основном проходит по границе (в том числе ребрам). Успешный алгоритм оказался гибридом методов наискорейшего и покоординатного спуска, его работа напоминает хорошо известный симплекс-метод [33]. Алгоритм можно описать следующим образом.

Поиск начинается с некоторой случайно выбранной внутренней точки z_{aj} . Затем производится некоторое число итераций — «раундов переговоров», в которых «участвуют» струи (j). Каждый «раунд» заключается в том, чтобы перебрать по очереди все частицы (a), и для каждой из них перераспределить z_{aj} так, чтобы уменьшить минимизируемую функцию Ω_R . «Переговоры» по каждой частице сводятся к вычислению градиента Ω_R по вектору переменных z_{aj} при фиксированном a , а перераспределение заключается в том, что z_{aj} сдвигается против градиента так далеко, насколько это возможно без выхода за границы области; если новое значение функции Ω_R не меньше старого, то алгоритм предусматривает уменьшение шага (хотя этого на прак-

тике никогда не требуется). Поиск заканчивается после «раунда», в котором ни одна частица не поменяла своего положения.

Важно повторить поиск несколько раз с разных начальных точек (выбираемых случайно), так как возможно наличие локальных минимумов. Однако структура функции такова, что глобальный минимум «притягивает» к себе случайную начальную точку в рамках данного алгоритма с наибольшей вероятностью, так что некоторые трудности могут возникнуть лишь при исчерпывающем поиске всех локальных минимумов, что редко нужно.

Имеются две важные оптимизации, весьма существенно ускоряющие алгоритм. Первая основана на наблюдении, что как функция Ω_R , так и ее градиенты не зависят от z_{aj} напрямую, а только через q_j , которые зависят от z_{aj} линейно. Поэтому после сдвига частицы достаточно «подправить» q_j , учтя изменения соответствующих z_{aj} . Вторая оптимизация связана с вычислением проекции градиента на границу (что происходит на всех шагах после первого «раунда»). Такая проекция определена неоднозначно, поэтому важно выбирать ее оптимальным образом (в самом деле, даже просто вычисление неоптимальной проекции нетривиально ввиду сложности области (25)). В работе [5] описан изящный цикл, который может быть реализован в виде быстрых побитных операций, производящих оптимальный выбор.

Как показали тесты на большом массиве реальных событий*, полученных на ускорителе LEP2, наша реализация лишь незначительно медленнее, чем стандартный рекомбинационный алгоритм k_T . При этом можно сделать вывод о том, что по стандартным тестам для АОС (таким как процент событий, для которых алгоритм правильно восстанавливает исходную конфигурацию партонов для событий, полученных с помощью стандартных генераторов) наш «оптимальный» АОС, как минимум, не хуже алгоритма k_T . Более того, имеются основания ожидать, что «оптимальный» АОС имеет существенно лучшее асимптотическое поведение при большом числе частиц (линейный рост вместо квадратичного для алгоритма k_T). Одно это обстоятельство (при прочих равных) может сделать его предпочтительным для использования в экспериментах на новых ускорителях, таких как LHC, где при больших энергиях в каждом соударении рождается больше частиц. Код на фортране, реализующий найденный алгоритм, доступен в Интернете [34].

8. НОВЫЕ ВОЗМОЖНОСТИ

То обстоятельство, что обсуждаемый АОС «погружен» в достаточно богатую теорию, открывает новые возможности, не известные или не существующие в контексте стандартных алгоритмов. Укажем на две из них.

*Тесты проведены Пабло Ашаром (Pablo.Achard@cern.ch) из коллаборации L3.

Наш АОС, наряду с 4-импульсами струй, автоматически дает дополнительную информацию о событии: значения величин Y_j для каждой струи, а также E_{soft} . Величина Y_j , естественно, интерпретируется как ширина струи (точнее, динамическая ширина, так как значения Y_j зависят от распределения энергии среди частиц, составляющих струю). При этом Y_j и E_{soft} определены модельно-независимым образом. Поскольку для соответствующего «жесткого» партонного события эти величины в ведущем приближении равны нулю (для струи, состоящей из одной частицы, $Y_j = 0$), то ненулевые значения этих параметров свидетельствуют о вкладе непертурбативных степенных или высших логарифмических поправок. Роль таких поправок вызывает теоретический интерес (см., например, [35]), и указанная модельно-независимая параметризация могла бы дать кинематическую основу для систематического изучения динамики таких поправок.

Вторая интересная новая возможность связана с отказом от широко используемой жесткой классификации и, соответственно, отбора событий по числу струй. Мы уже видели на примере поиска топ-кварка в чисто адронном канале распада, что регуляризованная форма такого отбора (приводящая к наблюдаемым, которые интерпретируются как нецелое число струй, см. (17)) может давать ощутимые преимущества. Рассмотрим любой процесс, в котором в низшем порядке теории возмущений в конечном состоянии должно быть N жестких партонов, и который изучается для измерения какого-либо параметра (скажем, массы какой-то нестабильной частицы). Общепринятая процедура состоит в том, чтобы отбирать события с N струями с помощью какого-либо АОС и затем строить наблюдаемые уже на этом подмножестве. Однако такой отбор эквивалентен использованию наблюдаемых, обращающихся в нуль на событиях с числом струй, не равным N , т. е. претерпевающих разрыв на границах между областями событий с числом струй, равным и не равным N . Ясно, что неоптимальность таких наблюдаемых (т. е. потери в отношении сигнал/шум) тем больше, чем значительнее роль высших теоретических поправок, ухудшающих адекватность простейшей физической картины. Рецепт регуляризации состоит в том, чтобы не просто отбирать события, а приписывать им определенный вес, и отбрасывать лишь события с достаточно малым весом.

Например, можно было бы поступить следующим (или подобным) образом. Сначала проведем параметризацию событий с помощью $y_N = \min \Omega_R$, фигурирующего в левой части критерия (30) и вычисленного для заданного N . Построим квазинепрерывную наблюдаемую η , соответствующую измеряемому параметру, как функцию y_N . Такая наблюдаемая будет убывать как при малых, так и при больших y_N (т. е., соответственно, в направлениях событий с числом струй меньше и больше N). Теперь события можно отбирать в зависимости не от произвольного y_{cut} , а от величины η и с учетом имеющихся вычислительных ресурсов. Ясно, что говорить об отборе событий с

фиксированным числом струй здесь уже нельзя, поскольку удержанные события с небольшим η могут соответствовать событиям, которые при обычном подходе были бы отброшены как имеющие слишком большое или, наоборот, малое число струй. Во всяком случае, ясно, что теория квазиоптимальных наблюдаемых открывает широкое поле деятельности для осмысленного и в значительной степени детерминированного построения процедур обработки данных (в отличие от общепринятого подхода, где такие процедуры строятся методом проб и ошибок более или менее «вслепую»), а найденный оптимальный АОС вписывается в рамки этой теории вполне естественно.

ЗАКЛЮЧЕНИЕ

Итак, найденное определение струй (30) является следствием достаточно глубокой теории, учитывающей такие разные аспекты проблемы измерения в случае многочастичных адронных состояний, как структура ошибок измерения многомодульных калориметрических детекторов, структура оптимальных наблюдаемых, структура квантово-полевых наблюдаемых. Найденный АОС имеет прозрачный физический смысл (искомая конфигурация струй минимизирует потерю физической информации при переходе к сокращенному описанию события в терминах небольшого числа струй), является оптимальным (т. е. минимизирует потерю информации наилучшим образом), лишен недостатков обычных алгоритмов (например, решается проблема множественных конфигураций струй), при этом совмещает их лучшие свойства (удобство с точки зрения теоретических расчетов), допускает эффективную и стабильную реализацию в виде компьютерной программы [34]. При этом найденный АОС естественно «погружен» в теорию квазиоптимальных наблюдаемых, что открывает новые возможности для построения более совершенных процедур обработки данных.

Удивительно то, что разные соображения (как чисто кинематические, исходящие из анализа ошибок измерения, так и более динамические, учитывающие квантово-полевые и хромодинамические аспекты) приводят к одному и тому же определению струй. Это, видимо, следует считать указанием на то, что найденный «оптимальный» АОС является в существенном единственным.

Автор признателен А. П. Исаеву за помощь и Д. Ю. Бардину за ценные замечания. Работа частично поддержана грантом РФФИ 99-02-18365.

СПИСОК ЛИТЕРАТУРЫ

1. Tkachov F. V. // Phys. Rev. Lett. 1994. V. 73. P. 2405; Phys. Rev. Lett. 1995. V. 74. P. 2618; hep-ph/9901332.

2. *Tkachov F. V.* // Int. J. Mod. Phys. A. 1997. V. 12. P. 5411; hep-ph/9601308.
3. *Sveshnikov N. A., Tkachov F. V.* // Phys. Lett. B. 1996. V. 382. P. 403; hep-ph/9512370.
4. *Tkachov F. V.* // IJMPA (in print); hep-ph/9901444 (revised Jan. 2000).
5. *Grigoriev D. Y., Tkachov F. V.* // Proc. of the XIV Intern. Workshop «High Energy Physics and Quantum Field Theory» (QFTHEP'99), Moscow, May 27–June 2 / Eds. B. B. Levchenko, V. I. Savrin, M., 1999; hep-ph/9912415.
6. *Ткачев Ф. В.* // Направлено в «Письма в ЭЧАЯ»; physics/0001019.
7. *Drell S. D., Levy D., Yan T.-M.* // Phys. Rev. 1969. V. 187. P. 2159;
Feynman R. P. Photon–Hadron Interactions. Benjamin. 1972.
8. *Hanson G. et al.* // Phys. Rev. Lett. 1975. V. 35. P. 1609.
9. *Ellis J., Gaillard M. K., Ross G. G.* // Nucl. Phys. B. 1976. V. 111. P. 253.
10. *Sterman G., Weinberg S.* // Phys. Rev. Lett. 1977. V. 39. P. 1436.
11. *Barlow R.* // Rep. Prog. Phys. 1993. V. 56. P. 1067.
12. *Mirkes E.* Theory of Jets in Deep Inelastic Scattering: Habilitation thesis. Karlsruhe Univ., 1997.
13. *Babcock J. B., Cutkosky R. E.* // Nucl. Phys. B. 1980. V. 176. P. 113.
14. *Brandt S. et al.* // Phys. Lett. 1964. V. 12. P. 57;
Farhi E. // Phys. Rev. Lett. 1977. V. 39. P. 1587.
15. *Sjöstrand T.* // Comp. Phys. Comm. 1983. V. 28. P. 229.
16. *Moretti S., Lonnblad L., Sjöstrand T.* // JHEP. 1998. V. 9808. P. 1; hep-ph/9804296.
17. *Bartel W. et al.* (JADE Collaboration) // Z. Phys. C. 1986. V. 33. P. 23.
18. *Bethke S. et al.* // Nucl. Phys. B. 1992. V. 370. P. 310.
19. *Catani S. et al.* // Phys. Lett. B. 1991. V. 269. P. 432.
20. *Youssef S.* // Comp. Phys. Comm. 1987. V. 45. P. 423.
21. *Dydak F.* Talk at the IX Intern. Workshop on High Energy Physics, Zvenigorod, Sept. 1994.
22. *Fox G. C., Wolfram S.* // Nucl. Phys. B. 1979. V. 149. P. 413.
23. *Georgi H., Machacek M.* // Phys. Rev. Lett. 1977. V. 39. P. 1237;
Farhi E. // Ibid. P. 1587;
De Rujula A. et al. // Nucl. Phys. B. 1978. V. 138. P. 387;
Pi S.-Y., Jaffe R. L., Low F. E. // Phys. Rev. Lett. 1978. V. 41. P. 142;
Fox G. C., Wolfram S. // Ibid. P. 1581;
Basham C. L. et al. // Ibid. P. 1585;
Brandt S., Dahmen H. D. // Z. Phys. C. 1979. V. 1. P. 61;
Donoghue J. F., Low F. E., Pi S.-Y. // Phys. Rev. D. 1979. V. 20. P. 2759.
24. Physics at LEP2 / Eds. G. Altarelli, T. Sjöstrand, F. Zwirner. CERN 96-01. V. 1.
25. *Bhat P. C., Prosper H., Snyder S. S.* // Intern. J. Mod. Phys. A. 1998. V. 13. P. 5113; hep-ex/9809011.
26. *Ткачев Ф. В.* Доклад на XV Междунар. совещ. QFTHEP'2000, Тверь, 14–20 сент. 2000 г.; hep-ph/0012210.
27. *Rudin W.* Functional Analysis. McGraw-Hill, 1973.
28. *Shifman M. A., Vainshtein A. I., Zakharov V. I.* // Nucl. Phys. B. 1979. V. 147. P. 385.
29. *Tkachov F. V.* // Phys. Lett. B. 1983. V. 125. P. 85; hep-ph/9901214.
30. *Tkachov F. V.* // Phys. Part. Nucl. 1994. V. 25. P. 649; hep-ph/9701272.

31. *Tkachov F. V.* // Phys. Lett. B. 1997. V. 412. P. 350; hep-ph/9703424.
32. *Eadie W. T. et al.* Statistical Methods in Experimental Physics. Amsterdam, 1971.
33. *Калиткин Н. Н.* Численные методы. М.: Наука, 1978.
34. <http://www.inr.ac.ru/ftkachov/projects/jets/>
35. *Dokshitzer Y. L., Webber B. R.* // Phys. Lett. B. 1995. V. 352. P. 451.