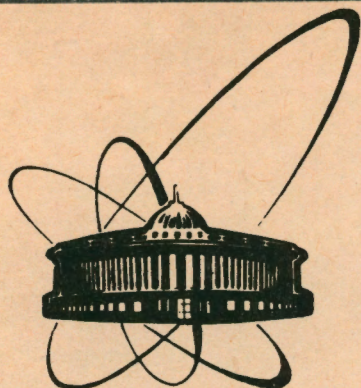


9-367



ОБЪЕДИНЕННЫЙ  
ИНСТИТУТ  
ЯДЕРНЫХ  
ИССЛЕДОВАНИЙ  
ДУБНА

P10-90-367

Г.А.Ососков, Н.А.Русакович, Н.И.Чернов

БЫСТРЫЙ АЛГОРИТМ  
АНАЛИЗА МНОГОТРЕКОВЫХ СОБЫТИЙ  
С ОБРАЗОВАНИЕМ КОРОТКОЖИВУЩИХ ЧАСТИЦ

Направлено в журнал "Nuclear Instruments  
and Methods"

1990

## 1. ВВЕДЕНИЕ

Характерной особенностью взаимодействия адронов при высоких энергиях является большая множественность частиц в конечном состоянии. В экспериментах на фиксированной мишени при энергиях свыше 1 ТэВ средняя множественность составляет десятки частиц. В этих условиях большое значение приобретает поиск алгоритмов, позволяющих произвести фильтрацию событий на уровне триггера по некоторым характерным признакам.

В экспериментах, нацеленных на изучение физики D- и В-мезонов, таким признаком является факт наличия вторичных вершин, происходящих от распада этих мезонов. Трудности в организации эффективного отбора событий с образованием мезонов связаны со следующими обстоятельствами:

- малая длина пробега до распада (при энергии мезона  $\sim 100$  ГэВ его средний пробег порядка 1 см);
- крайне низкий уровень отношения сигнал/фон ( $\sim 10^{-4} \div 10^{-5}$ ) из-за малого сечения образования мезонов;
- малые углы разлета вторичных частиц.

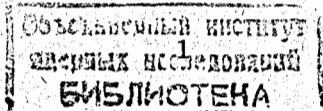
Мы рассмотрим алгоритм, позволяющий при относительно малом объеме вычислительных операций сделать вывод о наличии в событии, как минимум, одной вторичной вершины на некотором расстоянии от точки первичного взаимодействия. Предлагаемый алгоритм ориентирован на конкретную модель вершинного детектора, но эта ориентация проявляется только в выборе численных значений констант. Идея алгоритма довольно универсальна, в чем читатель легко сможет убедиться.

## 2. МОДЕЛЬ ВЕРШИННОГО ДЕТЕКТОРА

Вершинный детектор (см. рис. 1) представляет собой блок дрейфовых камер (ДК) с плоскостями, ортогональными пучку. Камеры для X- и Y-проекций чередуются друг с другом (ось Z направлена, как принято, вдоль пучка). Всего имеется  $10 \div 20$  ДК в каждой проекции. Расстояние от мишени до первой камеры — 1 м, длина блока камер — 2 м.

Параметры ДК:

- точность измерения координат — 0,1 мм,
- двухтрековое разрешение — 1 мм,



- эффективность —  $98 \div 99\%$ ,
  - уровень шумов — 1-2 шумовых сигнала на камеру.
- Магнитное поле отсутствует, поэтому все треки — прямолинейные.

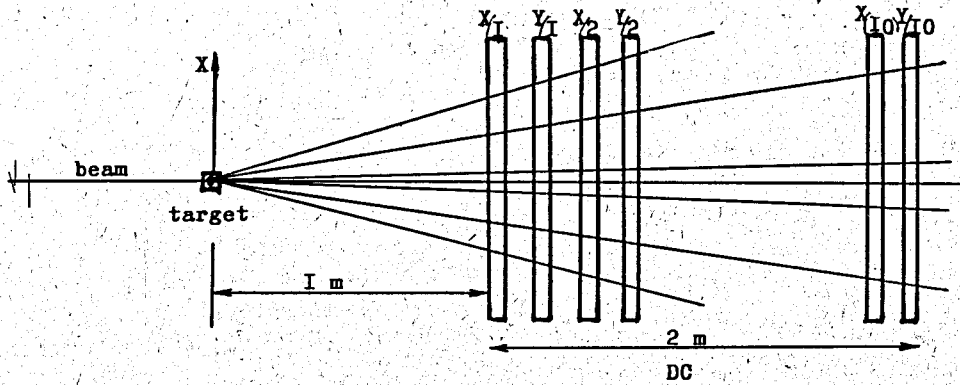


Рис. 1

### 3. МОДЕЛЬ СОБЫТИЯ

Будем считать, что в событии наблюдается в среднем 15 вторичных (заряженных) частиц, число которых распределено по закону Пуассона. Среди вторичных частиц появляются короткоживущие мезоны (D, B), являющиеся предметом исследования. Они быстро распадаются, порождая вторичные вершины (см. рис. 2).

Мезоны появляются всегда парами, т.е. два B-мезона разных зарядов или два D-мезона. Длина  $L_0$  пробега мезона до его распада распределена экспоненциально со средним  $L_0 \approx 1 \div 2$  см.

Таким образом, находясь так близко от первичной вершины, вторичные вершины оказываются почти что в самой мишени, создавая этим основную трудность выделения событий, содержащих мезоны.

Средняя множественность вторичных (заряженных) частиц после распада мезона равна  $4 \div 5$ . Число вторичных частиц распределено по закону Пуассона.

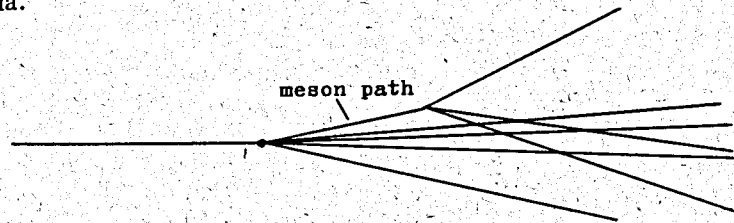


Рис. 2

Все вторичные частицы разлетаются под небольшими углами. Функцию распределения угла разлета обозначим  $F_\alpha(x)$ . В качестве грубого приближения зададим плотность  $F_\alpha(x)$  в виде треугольной функции, изображенной на рис. 3 (здесь  $\alpha_{\max} = 30^\circ \div 40^\circ$ ).

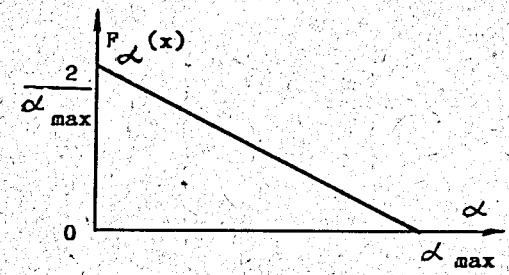


Рис. 3

### 4. ТРЕКИНГ

Обработка данных начинается с выделения треков. На каждой проекции X и Y это делается независимо.

Для грубой (с ошибкой  $5 \div 10$  мм) аппроксимации прямолинейного трека достаточно соединить одну его точку в любой из камер с центром мишени. Далее следует собрать все точки в других камерах, лежащие в окрестности полученной прямой линии, и подогнать к ним прямую взвешенным методом наименьших квадратов (МНК).

Замечание 1. Треки следует искать вне некоторого конуса вокруг направления пучка, т.к. треков с малым углом вылета много, они сливаются друг с другом и малоинформативны.

Замечание 2. Для обеспечения большей точности поиска следует начинать построение трека с последней камеры и заканчивать на первой.

После отслеживания N треков получаем для каждого i-го трека  $n_i$  точек с координатами  $(x_1, z_1), \dots, (x_{n_i}, z_{n_i})$  (для краткости мы ограничимся проекцией XZ, все последующие рассуждения справедливы и для плоскости YZ). Обозначим для трека с номером i ( $i = 1, N$ )

$$\begin{aligned}
 S_X^{(i)} &= x_1 w_1 + x_2 w_2 + \dots + x_{n_i} w_{n_i}, \\
 S_{XZ}^{(i)} &= x_1 z_1 w_1 + \dots + x_{n_i} z_{n_i} w_{n_i}, \\
 S_Z^{(i)} &= z_1 w_1 + \dots + z_{n_i} w_{n_i}, \\
 S_{ZZ}^{(i)} &= z_1^2 w_1 + \dots + z_{n_i}^2 w_{n_i}, \\
 S_0^{(i)} &= w_1 + \dots + w_{n_i}.
 \end{aligned} \tag{1}$$

Здесь веса полагаются равными  $w_i = 1/D(x_i)$ , где  $D(x_i)$  — дисперсия измерения координаты  $x_i$ , складывающаяся из случайной ошибки из-

мерения в камере и из случайного отклонения траектории частицы от прямой линии под действием кулоновского рассеяния и т.д. (по этой причине дисперсия  $D(x_i)$  минимальна на первой камере  $(-0,01)$  и возрастает к последней камере).

Уравнение трека  $x = a \cdot z + b$ , оцененное по МНК, имеет параметры

$$a = \frac{S_{XZ} S_0 - S_X S_Z}{S_{ZZ} S_0 - (S_Z)^2}, \quad (2)$$

$$b = \frac{S_{ZZ} S_X - S_{XZ} S_Z}{S_{ZZ} S_0 - (S_Z)^2}.$$

## 5. ПОИСК ВЕРШИНЫ

Статистическая проверка гипотезы о том, что все треки сходятся в одной общей вершине, выполняется в два этапа: а) оценка положения вершины, б) пороговая процедура проверки по специальному критерию (goodness-of-fit test). Рассмотрим их по очереди.

5.1. Оценка координат вершины. Пусть  $(x_0, z_0)$  — гипотетические координаты вершины. Если гипотеза верна, то отклонение

$$d_i = x_0 - a_i z_0 - b_i$$

$i$ -го трека от вершины распределено по нормальному закону  $N(0, D_i)$ , где дисперсию  $D_i$  легко вычислить:

$$D_i = \text{Var}(x_0 - a_i z_0 - b_i) = \frac{S_0^{(i)} z_0^2 - 2S_Z^{(i)} z_0 + S_{ZZ}^{(i)}}{S_{ZZ}^{(i)} S_0^{(i)} - (S_Z^{(i)})^2},$$

Если поместить начало координат в центр мишени, то координаты вершины  $(x, z)$  будут близки к нулю и можно воспользоваться приближением

$$D_i \approx \frac{S_{ZZ}^{(i)}}{S_{ZZ}^{(i)} S_0^{(i)} - (S_Z^{(i)})^2}, \quad i = \overline{1, N}. \quad (3)$$

Величины  $D_i$  в (3) уже не зависят от  $z_0$ . Обозначим  $c_i = 1/D_i$ .

Заметим, что величины  $D_i$  зависят только от координат дрейфовых камер, поэтому их значения мало отличаются для разных треков, что составляет возможность для еще более грубой аппроксимации  $D_i = D_0 = \text{const}$ .

Поскольку все треки независимы друг от друга, то для оценки положения вершины следует минимизировать величину

$$\sum_{i=1}^N c_i (x_0 - a_i z_0 - b_i)^2.$$

Минимизация с помощью МНК дает оценки

$$x_0 = \frac{(\sum a_i^2 c_i) \cdot (\sum b_i c_i) - (\sum a_i c_i) \cdot (\sum a_i b_i c_i)}{(\sum c_i) (\sum a_i^2 c_i) - (\sum a_i c_i)^2}, \quad (4)$$

$$z_0 = \frac{(\sum a_i c_i) \cdot (\sum b_i c_i) - (\sum c_i) \cdot (\sum a_i b_i c_i)}{(\sum c_i) \cdot (\sum a_i^2 c_i) - (\sum a_i c_i)^2}. \quad (5)$$

5.2. Проверка одновершинности. Если гипотеза верна, то индивидуальные отклонения треков от вершины  $d_i = x_0 - a_i z_0 - b_i$  после нормировки на  $\sqrt{D_i}$  будут распределены по нормальному закону  $N(0, 1)$ . В силу независимости одновременная малость величин  $d_i/\sqrt{D_i}$  является вполне естественным критерием для принятия гипотезы одновершинности. Наоборот, если хотя бы одна из них велика, гипотеза отвергается (такие методы называются "промах", т.к. мы ищем трек, "пролетающий" мимо общей вершины).

Для строгой формулировки этой идеи рассмотрим статистику

$$d_{\max}^{(N)} = \max_{1 \leq i \leq N} \{|d_i|/\sqrt{D_i}\}. \quad (6)$$

Если гипотеза одновершинности справедлива, то  $d_{\max}^{(N)}$  имеет распределение

$$F_N(x) = \text{Prob}\{d_{\max}^{(N)} < x\} = \left\{ \frac{1}{\sqrt{2\pi}} \int_{-x}^x \exp(-t^2/2) dt \right\}^N.$$

Распределение  $F_N(x)$  легко табулируется. По соответствующим таблицам можно найти величину

$$P = 1 - F_N(d_{\max}^{(n)}) \quad (7)$$

и считать ее уровнем доверия для гипотезы одновершинности: когда  $P$  велико ( $P > P_{\text{пор}}$ ), гипотеза принимается, а когда  $P < P_{\text{пор}}$ , то отвергается. По величине  $P_{\text{пор}}$  легко вычислить значение  $d_{\text{пор}}^{(N)}$  (по формуле (7)) и браковать гипотезу в случае  $d_{\max}^{(N)} > d_{\text{пор}}^{(N)}$ , что значительно удобнее в вычислительных процедурах.

## 6. ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМА

Известно, что проверка гипотезы может приводить к ошибкам двух видов:

Ошибка 1-го рода: гипотеза  $H_0$  верна, но проверка ее отвергает;

Ошибка 2-го рода: гипотеза  $H_0$  неверна, но проверка ее принимает. В нашем случае  $H_0$  — гипотеза одновершинности события и вероятность ошибки 1-го рода в точности равна  $P_{\text{пор}}$  из п.5.2. Эту вероятность можно сделать сколь угодно близкой к нулю путем небольшого изменения пороговой величины  $d_{\text{пор}}^{(N)}$ .

Вероятность ошибки 2-го рода от величины  $d_{\text{пор}}^{(N)}$  зависит слабо, но эта вероятность является более важной для экспериментатора, т.к. она определяет процент потерянных многовершинных (полезных) событий. Для оценки этой вероятности предлагается следующий метод.

Будем считать, что ошибка 2-го рода происходит, если все треки, выходящие из вторичных вершин (точнее — их продолжения назад, за эти вершины), проходят на расстоянии  $|d_i| < c\sqrt{D_i}$  от первичной вершины (здесь  $c$  — постоянная, равная 4 или 5, а остальные обозначения введены в п.5). Будем называть это условием неотличимости для вторичных треков (имея в виду неотличимость принадлежности трека к первичной или ко вторичной вершине).

Если  $k$  — число вторичных вершин, то вероятность ошибки 2-го рода  $P_2 = P_B^k$ , где  $P_B$  — вероятность неотличимости всех треков какой-либо одной вторичной вершины. Далее, если число  $N$  треков, выходящих из вторичной вершины, распределено по закону Пуассона со средним  $\lambda$ , то

$$P_B = \sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \exp(-\lambda) \cdot P_T^N = \exp(-\lambda(1 - P_T))$$

где  $P_T$  — вероятность неотличимости одного конкретного трека. Поэтому

$$P_2 = \exp(-k \cdot \lambda \cdot (1 - P_T)) \quad (8)$$

(величину  $\lambda$  следует брать несколько заниженной, т.к. часть треков теряется из-за неэффективности камер или попадает в узкий конус, который программа не анализирует — см. п. 4).

Если  $L$  — длина пробега соответствующего мезона до распада (или иначе — расстояние от первичной вершины до вторичной), а  $\alpha$  — угол вылета рассматриваемого трека из вторичной вершины, то условие его неотличимости выполняется при  $\alpha L < c\sqrt{D_1}$ . Величины  $D_1$  можно грубо считать одинаковыми для всех треков (см. п.5.1) и тогда  $D_1 \approx D_0$ . Величины  $\alpha$  и  $L$  случайны и независимы, их распределения описаны в п.3. Используя это, получаем вероятность условия неотличимости для одного трека

$$P_T = \text{Prob}\{\alpha L < c\sqrt{D_0}\} = \frac{1}{L_0} \int_0^{\infty} e^{-L/L_0} F_{\alpha} \left( \frac{c\sqrt{D_0}}{L} \right) dL \quad (9)$$

Подставляя в (9) конкретный вид  $F_{\alpha}(x)$ , получаем  $P_T$ , а затем из формулы (8) — вероятность ошибки второго рода  $P_2$ .

Для модели, описанной в п.п.2, 3, мы провели численный расчет величины  $P$  (при значениях  $k = 2$ ,  $\lambda = 3$ ,  $\alpha_{\max} = 35^\circ$ ,  $c = 5$ ,  $\sqrt{D_0} = 0,1$  мм). Результаты представлены на рис. 4 в виде зависимости  $P_2$  от средней длины пробега мезонов ( $L_0$ ).

## 7. ЧИСЛЕННАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМА

В соответствии с п.п.2, 3 на персональном компьютере типа IBM PC была сгенерирована статистика с несколько тысяч событий, достаточная для численных оценок ошибок первого и второго рода. Результаты численного эксперимента полностью подтвердили расчеты, сделанные в п.6. Вероятность  $P_{\text{пор}}$ , заданная в пороговой процедуре, соответствовала

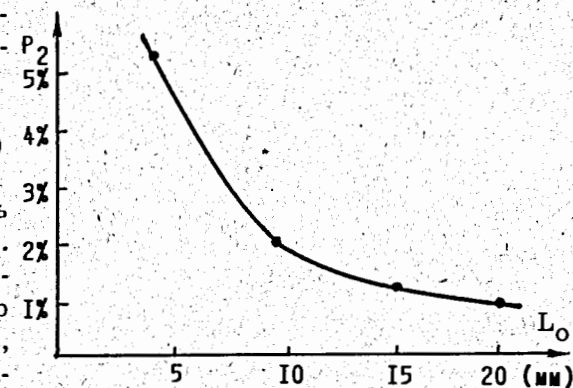


Рис. 4

ошибке 1-го рода, а вероятность ошибки 2-го рода слабо зависела от  $P_{пор}$  (даже при значениях  $P_{пор}$  порядка  $10^{-10}$ ) и приближенно могла быть рассчитана по формулам (8-9). В нашей конкретной модели  $P_2 \approx 2\%$ .

Дополнительно в численном эксперименте было сгенерировано несколько тысяч фоновых событий, происходящих от взаимодействия пучка вне мишени. В этом случае предположения о малости  $z_0$ , сделанные при выводе формулы (3), нарушаются. Тем не менее, оценки (4) и (5) определяют положение вершины достаточно точно, и простая проверка

$$|z_0| < 3 \text{ см}$$

позволяет отсеять 99,9% фоновых событий.

## 8. ЗАКЛЮЧЕНИЕ

Отметим следующие возможности программно-аппаратной реализации метода:

— в формулах (1-6) величины  $S_Z^{(i)}$ ,  $S_{ZZ}^{(i)}$  не зависят от положений треков, а определяются номерами сработавших камер, т.к. Z-координаты камер фиксированы; поэтому данные величины могут быть просто затабулированы; для более грубой аппроксимации можно считать их вообще постоянными для всех треков;

— подсчет величин  $S_X^{(i)}$ ,  $S_{XZ}^{(i)}$  и сумм, входящих в формулы (4,5), требует ок. 50 арифметических операций на один трек и может быть реализован на процессорах параллельного действия;

Таким образом, в принятых моделях события и вершинного детектора результаты работы предложенного алгоритма оказались следующими:

- выбирается около 98% полезных (многовершинных) событий;
- ошибочно выбирается не более 1% одновершинных событий, происходящих в мишени;
- из фоновых событий, происходящих от взаимодействий пучка вне мишени, ошибочно выбирается не более 0,1%.

Рукопись поступила в издательский отдел  
30 мая 1990 года.

Ососков Г.А., Русакович Н.А., Чернов Н.И. P10-90-367  
Быстрый алгоритм анализа многотрековых событий  
с образованием короткоживущих частиц

Для модели вершинного детектора (ВД), находящегося вне магнитного поля в эксперименте с фиксированной мишенью, когда пробеги короткоживущих вторичных частиц сравнимы с размерами мишени, а их разлет ограничен узким конусом ( $< 30^\circ$ ), предложен алгоритм быстрой фильтрации одновершинных событий. Алгоритм основан на пороговой процедуре, требующей не более 50 арифметических операций на 1 трек, что делает его пригодным для применения в триггере высокого уровня. Теоретические и численные расчеты эффективности алгоритма для модельного ВД дают вероятность потери многовершинного события около 2%, а вероятность ошибочного пропуска одновершинного события — менее 1%.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1990

Перевод авторов

Ososkov G.A., Russakovich N.A., Chernov B.I. P10-90-367  
Fast Algorithm for Analysis of Multitrack Events  
with Production of Short-Lived Particles

An algorithm of fast filtration of single-vertex events is proposed for a model of the vertex detector (VD) situated outside the magnetic field in the fixed target experiments with a short predecay path of short-lived particles (comparable with the target size) and small exit angles of secondary particles ( $< 30^\circ$ ). The algorithm is based on a threshold procedure requiring no more than 50 floating point operations per track that makes it applicable on a trigger of high level. Theoretical and numerical computations of the algorithm efficiency for a VD version shown that the probability of a multi-vertex event loss is about 2% and the probability of a single-vertex event erroneous selection is less than 1%.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1990