

**Д.М. Подорожный, Е.Б. Постников, Л.Г. Свешникова,  
А.Н. Турундаевский**

**ПРИМЕНЕНИЕ МНОГОМЕРНЫХ СТАТИСТИЧЕСКИХ  
МЕТОДОВ ДЛЯ ИНТЕРПРЕТАЦИИ ДАННЫХ ИЗМЕРЕНИЙ  
МНОГОКАНАЛЬНОЙ АППАРАТУРЫ  
НА ПРИМЕРЕ СПЕКТРОМЕТРА КЛЕМ**

Препринт НИИЯФ МГУ – 2003-12/725

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
им. М.В. ЛОМОНОСОВА

---

НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ  
ЯДЕРНОЙ ФИЗИКИ им. Д.В. СКОБЕЛЬЦЫНА

**Д.М. ПОДОРОЖНЫЙ, Е.Б. ПОСТНИКОВ, Л.Г. СВЕШНИКОВА,  
А.Н. ТУРУНДАЕВСКИЙ**

**ПРИМЕНЕНИЕ МНОГОМЕРНЫХ СТАТИСТИЧЕСКИХ  
МЕТОДОВ ДЛЯ ИНТЕРПРЕТАЦИИ ДАННЫХ ИЗМЕРЕНИЙ  
МНОГОКАНАЛЬНОЙ АППАРАТУРЫ  
НА ПРИМЕРЕ СПЕКТРОМЕТРА КЛЕМ**

Препринт НИИЯФ МГУ – 2003-12/725

УДК 520.626:524.1

ББК 22.38

П76

**D.M. Podorozhnyi, E.B. Postnikov, L.G. Sveshnikova, A.N. Turundaevsky**

E-mail: postn@rbcmail.ru

**THE MULTIVARIATE STATISTICAL TECHNIQUE FOR MULTI-CHANNEL APPARATUS  
ON EXAMPLE OF KLEM-SPECTROMETER**

Preprint of SINP MSU - 2003-12/725

Abstract

This paper is about the multivariate statistical technique solving the problems of a certain kind: when we are unable to measure the interesting physical parameters directly, but we obtain some information about them and this information is multidimensional. The problems of such a kind appear in the CR physics if we use for the energy and another primaries characteristics determination such detectors, which allow to measure a large number of parameters depending on these characteristics.

In the network of the multivariate technique the algorithms of the primary energy and the primary CR spectral parameter  $\gamma$  estimation were developed. They have been investigated on example of the concrete gaging equipment – the KLEM detector, NUCLEON project. The results of computer experiments modeling the multivariate technique work are presented. We have also performed a comparison between the suggested approach and the one-dimensional one. The comparison results demonstrate the advantages of the new method in both the accuracy and the universality as applied to practical realization.

**Д.М. Подорожный, Е.Б. Постников, Л.Г. Свешникова, А.Н. Турундаевский**

**ПРИМЕНЕНИЕ МНОГОМЕРНЫХ СТАТИСТИЧЕСКИХ МЕТОДОВ  
ДЛЯ ИНТЕРПРЕТАЦИИ ДАННЫХ ИЗМЕРЕНИЙ  
МНОГОКАНАЛЬНОЙ АППАРАТУРЫ НА ПРИМЕРЕ СПЕКТРОМЕТРА KLEM**

Препринт НИИЯФ МГУ - 2003-12/725

Аннотация

Работа посвящена многомерной статистической методике решения задач определенного вида – когда интересующие нас физические параметры недоступны непосредственному измерению, а имеющаяся о них информация многомерна. Такие задачи возникают в физике космических лучей, если мы используем для определения первичной энергии и других характеристик первичных частиц детекторы, позволяющие измерять большое количество зависящих от этих характеристик параметров.

В рамках многомерной методики разработаны алгоритмы оценивания первичной энергии и показателя степенного спектра первичных КЛ. Они исследованы на примере конкретной измерительной аппаратуры – спектрометра KLEM, проект NUCLEON. В работе приведены результаты компьютерных экспериментов, моделирующих работу многомерной методики для этой аппаратуры. Проведено также сравнение предлагаемого подхода с одномерным, результаты которого демонстрируют преимущества новой методики как в точности, так и в универсальности с точки зрения практической реализации.

© Д.М. Подорожный, Е.Б. Постников, А.Н. Турундаевский, 2003

© НИИЯФ МГУ, 2003

## Введение

Ситуация, когда измерительная аппаратура позволяет одновременно регистрировать значения большого количества параметров, в физике космических лучей нередка. Например, микростриповый кремниевый детектор, использующийся в составе спектрометра KLEM в проекте NUCLEON [1] для регистрации углового распределения вторичных частиц, позволяет измерить значения нескольких сотен параметров (электрический сигнал, пропорциональный величине ионизационных потерь в стрипах детектора), каждый из которых несёт информацию об энергии первичной частицы.

Естественно, следует стремиться максимально учесть всю собираемую в эксперименте информацию, которая в рассматриваемом нами случае многомерна. Хорошо известно [2], что учет при обработке экспериментальных данных любого дополнительного измерения уменьшает (в худшем случае не увеличивает) погрешность интерпретации. Поэтому, если получаемые в процессе работы детектора многопараметрические данные предполагается обрабатывать в рамках стохастической, а не детерминистической, модели эксперимента, наибольшую точность при анализе банка данных будут обеспечивать многомерные, а не одномерные, статистические методы.

В данной работе мы рассмотрим два типа задач, которые могут быть решены оптимально, т.е. с наивысшей в рамках широкого класса алгоритмов точностью. Это задачи оценивания по многомерным данным многоканальных регистраторов, во-первых, одной или нескольких недоступных непосредственному измерению физических величин [3], например, первичной энергии, заряда и т.п., и во-вторых, показателя степенного первичного спектра (как для «неограниченного» спектра на полуинтервале, так и для участка спектра от  $E_1$  до  $E_2$ ).

Их решения были получены, а также создано реализующее математическую методику программное обеспечение и проведены тестовые компьютерные эксперименты, для конкретного измерительного прибора – спектрометра KLEM в составе аппаратуры NUCLEON. Тем не менее, аналогичное решение допускают задачи такого типа для любого детектора или измерительной процедуры, которая характеризуется большим числом регистрируемых параметров. Кроме того, основной блок рассматриваемой методики не имеет привязки к конкретному прибору, и в качестве входных данных для работы программы многомерной стат.обработки могут также быть использованы результаты измерений самими разными детекторами независимо от их структуры и принципов функционирования.

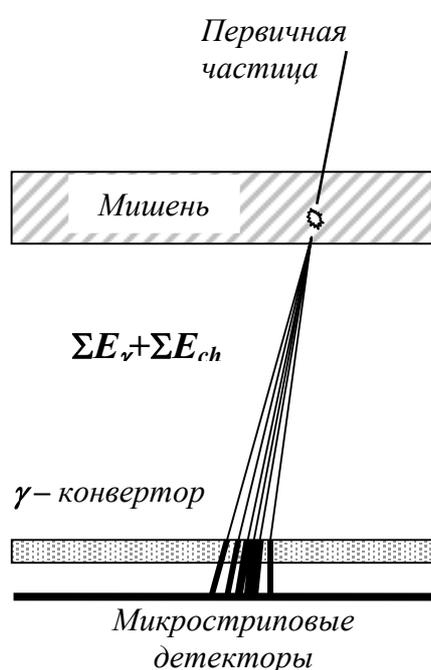
## Измерительная аппаратура KLEM

Задачей проекта NUCLEON явилась необходимость создания регистрирующей аппаратуры относительно небольшого веса при значительной величине светосилы с перспективой длительного времени экспозиции для проведения исследований КЛ (протонов и ядер) в широком (несколько порядков) энергетическом диапазоне единой методикой. Для этой цели был предложен и разработан новый подход к построению детектора энергии первичных частиц КЛ –

методика KLEM [4]. Суть данного подхода состоит в том, что энергия первичной частицы определяется из пространственной плотности распределения потока вторичных частиц, рожденных в тонкой мишени в первом акте неупругого взаимодействия и размножением в сверхтонкой толчковой установке – свинцовом (вольфрамовом) конверторе.

Схема спектрометра, на основе которой проводились вычислительные эксперименты, представлена на рис. 1. Аппарат состоит из мишени ( $\approx 20$  г/см<sup>2</sup> графита) и тонкого слоя свинца (10–20 г/см<sup>2</sup>), играющего роль конвертора гамма-квантов. Под конвертором расположены два перпендикулярно друг по отношению к другу ориентированных силиконовых микростриповых детектора. Величина шага стрипа варьировалась в пределах от 50 до 150 мкм.

**Рис. 1.** Схема измерительной аппаратуры KLEM



Первичная частица взаимодействует в мишени, где образуются вторичные гамма-кванты с энергией  $E_{\gamma}$  и заряженные частицы с энергией  $E_{ch}$ . Свинцовый конвертор нейтральных частиц, расположенный непосредственно перед регистрирующей плоскостью, конвертирует практически все вторичные гамма-кванты, падающие на этот слой вещества, в узкий пучок электронов за счет каскадного размножения в свинце.

Микростриповые детекторы состоят из большого количества узких длинных полосок (стрипов), в каждом из которых регистрируется сигнал, пропорциональный величине ионизационных потерь, вызванных попаданием в площадь стрипа вторичных частиц, и, следовательно, пропорциональный количеству попавших частиц.

Кроме того, перед мишенью располагается детектор заряда.

### Метод линейного наилучшего в среднеквадратичном (с.к.) оценивания случайных векторов

Простейший многомерный метод позволяет получить такое решение задачи оценивания одного или нескольких недоступных непосредственному измерению физических параметров, например, энергии первичной частицы, заряда и т.п., которое будет наиболее точным по сравнению с любым другим линейным алгоритмом их определения [2]. Например, для микрострипового детектора это утверждение означает, что какие бы коэффициенты при измеряемых детектором переменных  $N_i$  (величина сигнала в  $i$ -ом стрипе) мы не подобрали в эмпирической или умозрительной формуле для оценивания неизвестной первичной энергии

$$E = \sum_i C_i N_i + C_0,$$

значение оценки  $E$  все равно будет хуже, чем полученное в результате применения многомерной методики.

Несмотря на то, что метод линеен, «плата» за линейность вполне компенсируется его многомерностью: каждая из измеренных в эксперименте физических величин учитывается хотя и только в рамках линейной зависимости, но наилучшим образом, с учетом многочисленных связей измеряемых параметров как с неизвестной величиной, так и друг с другом. Поэтому на практике данный алгоритм работает гораздо лучше, чем любая «упрощенная» методика обработки данных, при которой все измеряемые дорогостоящей аппаратурой параметры заменяются одним, представляющим собой какую-либо их комбинацию, после чего на обычной двумерной плоскости строится нелинейная зависимость – степенная, полиномиальная и т.д. – неизвестной величины от этого параметра.\*

Кроме того, и в рамках многомерной идеологии существуют различные способы учета нелинейности, если она существенна в задаче из физических соображений.

Для реализации данного метода энергия  $E$  трактуется как случайная величина (т.е. одномерный случайный вектор), а все  $m$  измеряемых параметров (например, сигнал со стрипов детектора) – координаты  $m$ -мерного случайного вектора (обозначим его через  $\xi$ ). Суть метода состоит в формировании для случайной величины  $E$  линейной оценки, использующей всю доступную для измерения информацию:

$$E_{\text{оцен.}} = \sum_{i=1}^m b_i \xi_i + c, \quad (2)$$

где постоянные коэффициенты  $b_i$  и  $c$  выбираются таким образом, чтобы обеспечить минимум среднеквадратичной погрешности оценивания, т.е. среднего квадрата отклонения оценки  $E$  от ее истинного значения:

$$\mathbf{M}(E_{\text{оцен.}} - E)^2 = \left( \sum_{i=1}^m b_i \xi_i + c - E \right)^2 \sim \min \{ \text{по сравнению с любой другой линейной оценкой } E \}. \quad (3)$$

Искомые значения входящих в формулу для оценивания  $E$  коэффициентов даются приведенной в [2] теоремой. Для того, чтобы определить все неизвестные коэффициенты в (2) и тем самым полностью сформировать алгоритм оценивания первичной энергии, достаточно той априорной информации, которая содержится в достаточно большом объеме данных реальных измерений исследуемым детектором, либо модельных данных в том случае, если компьютерная модель адекватно описывает работу прибора. Этот банк полученных детектором в ходе регистрации КЛ данных (состоящий в примере с детектором KLEM из различных значений  $E$

---

\* Например, в разработанном нами ранее методе определения первичной энергии с помощью спектрометра KLEM вместо нескольких сотен или даже тысяч подлежащих измерению параметров  $N_i$  используется всего один параметр [1]

$$S = \sum_i \{ \ln^2(2r_i/H) \} N_i, \quad (1)$$

где  $r_i$  – расстояние  $i$ -го стрипа, в котором зарегистрировано  $N_i$  частиц, до оси ливня,  $H$  – расстояние до середины мишени, то есть до точки взаимодействия.

попавших на вход прибора первичных частиц и индуцированных ими значений  $N_i$ ) играет роль обучающей выборки.

Стандартный комплекс программного обеспечения GEANT давно успешно используется для моделирования ядернофизических процессов, лежащих в основе принципа действия детектирующей аппаратуры KLEM, в частности, он использовался для компьютерного тестирования и уточнения конструкции прибора с самого момента разработки проекта, поэтому для построения методики оценивания первичной энергии по многомерной методике также было применено компьютерное моделирование с помощью вышеназванного программного комплекса.

## **1. ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ ЭНЕРГИИ В КАЖДОМ ИНДИВИДУАЛЬНОМ СОБЫТИИ**

Таким образом, поэтапно предлагаемую методику определения энергии первичной частицы можно представить следующим образом:

1. Моделирование отклика прибора на прохождение через измерительную аппаратуру пучка первичных частиц определенного энергетического спектра, то есть формирование обучающей выборки;
2. Определение по соответствующей методике (описанной, например, в [2]) неизвестных констант в формуле (2) оценивания первичной энергии.

Теперь алгоритм оценивания уже построен, и его качество можно проверять как на модельных, так и на экспериментальных данных. Кроме того, в методику можно ввести один или несколько дополнительных параметров, описывающих наши априорные представления о характере статистической связи (точнее, о присутствующей в этой связи нелинейности, поскольку все линейные корреляции оцениваются автоматически) между измеряемыми и(или) оцениваемыми переменными. Сами параметры могут выбираться из физических соображений, а их оптимальные значения – по результатам вычислительного (компьютерного) эксперимента по моделированию работы сконструированной методики (подробнее об этом ниже). Естественно, следует брать как окончательные такие значения всех выбранных исследователем параметров, при которых качество оценивания интересующей его динамической величины – в данном случае первичной энергии – будет наивысшим. В рамках нашей методики «наивысшее качество» означает «минимальную среднеквадратичную погрешность оценивания энергии», см. (3), поскольку именно этому условию разрабатываемый алгоритм и должен удовлетворять.

Таким образом, следующий этап

3. Формирование еще одной случайной выборки – контрольной – путем моделирования работы измерительного прибора, и компьютерное тестирование методики оценивания на этой новой выборке. Вычисление погрешности оценивания (3) и определение оптимальных значений всех неизвестных параметров метода как доставляющих минимум погрешности оценивания.

Что касается формы энергетического спектра обучающей и контрольной выборок, то она должна определяться спецификой конкретной прикладной задачи, для решения которой нам необходим алгоритм оценивания энергии первичных частиц. Точность оценивания будет тем выше, чем больше априорной информации различного рода об исследуемом нами физическом процессе будет учтено. Так, если мы хотим потребовать от разрабатываемой методики, чтобы она минимизировала погрешность оценивания, среднюю по имеющему степенной характер реальному спектру КЛ, то следует в качестве обучающей и контрольной выборок «подавать» на вход прибора пучки первичных частиц, случайным образом выбранные из степенного распределения по энергии. Однако в степенном спектре преобладают частицы низких энергий, они же будут вносить наибольший вклад в среднеквадратичную погрешность, поэтому может представлять интерес и другая задача, а именно, если для нас более ценна точность восстановления энергии в широком диапазоне по логарифмической шкале (эта же ситуация возникает в случае, когда мы, как это делалось в проекте NUCLEON ранее, хотим вычислить погрешность энергетических измерений на монохроматических по  $E$  пучках первичных частиц, значения энергии которых равномерно расположены в логарифмической шкале:  $E = 10^{11}$ ,  $E = 10^{12}$ ,  $E = 10^{13}$  эВ и т.д.). Для решения данной задачи, естественно, следует брать в качестве обучающей выборку из такого энергетического спектра, в котором логарифм энергии распределен равномерно на всем интересующем интервале, а контрольной выборкой можно сделать стандартные энергетические монопучки.

Еще одно важное замечание заключается в том, что в качестве критерия, которому должна удовлетворять разрабатываемая методика, можно использовать не только условие (3) минимизации абсолютной погрешности оценивания энергии  $\mathbf{M}(E_{\text{оцен.}} - E)^2$ , но также можно модифицировать алгоритм таким образом, чтобы он доставлял минимум безразмерной *относительной погрешности*  $\mathbf{M}((E_{\text{оцен.}} - E) / E)^2$ , которая имеет более наглядный смысл. Для того, чтобы найти оценку первичной энергии наименьшей относительной погрешности, мы ищем ее в точно таком же виде, как и раньше (2):  $E_{\text{оцен.}} = \sum_{i=1}^m b_i \xi_i + c$ , но условие (3) в этом случае переписется в виде:

$$\mathbf{M}((E_{\text{оцен.}} - E) / E)^2 = \left( \sum_{i=1}^m b_i \xi_i + c - E \right)^2 \sim \min \{ \text{по сравнению с любой другой линейной оценкой } E \}. \quad (4)$$

После несложных преобразований данная задача сводится к предыдущей, решение которой приведено в [2]. Хотя выражение для оценки  $E$ , минимизирующей относительную погрешность  $\mathbf{M}((E_{\text{оцен.}} - E) / E)^2$ , в работе [2] не приводится, мы не приводим его и здесь в силу громоздкости, отметив лишь, что никакие изменения в алгоритме 1.–3. методики определения энергии первичной частицы не происходят, за исключением формулы, по которой на основе смоделированной обучающей выборки определяются коэффициенты  $b_i$  и  $c$  в выражении для  $E_{\text{оцен.}}$ .

Наконец, несколько слов о параметрах, позволяющих в рамках выбранной методики учесть нелинейность рассматриваемых физических процессов. Во-первых, еще в ходе компьютерных экспериментов с одномерным алгоритмом оценивания первичной энергии, использующим искусственный параметр  $S$  (1), было установлено, что наиболее сильная корреляция (т.е. линейная статистическая зависимость) наблюдается между  $E$  и  $N_i^a$ , где  $a \approx 1,2-1,4$ , а не 1. Этот факт вполне может быть учтен в многомерном алгоритме оценивания, точно так же, как и в случае, если наиболее сильная корреляция имела бы место между неизвестным параметром ( $E$ ) и любой известной функцией от измеряемых параметров, сколь сложна и нелинейна она бы не была. Для учета этой априорной информации достаточно просто на начальном этапе соответствующим образом трансформировать исходный банк данных и в дальнейшем работать с данными не по самим зарегистрированным сигналам  $N_i$ , а по  $N_i^a$ .

Кроме того, еще один фактор, позволяющий эффективно учитывать нелинейность лежащих в основе работы измерительной аппаратуры физических процессов, заложен внутри вычислительной процедуры многомерной оценивающей методики. Не вдаваясь в излишние тонкости и детали, отметим, что реализованный нами алгоритм включает в себя своеобразную «фильтрацию» мелких, несущественных и ложных взаимосвязей, возникших только благодаря конечности статистики, а также решает проблему некорректности вычислений.

## **АЛГОРИТМ ОЦЕНИВАНИЯ ПЕРВИЧНОЙ ЭНЕРГИИ ПО ДАННЫМ ИЗМЕРИТЕЛЬНОЙ АППАРАТУРЫ**

Подытоживая всё вышесказанное, сформулируем алгоритм оценивания первичной энергии уже применительно к конкретному детектору KLEM:

1. Определиться с тем, частицы какого энергетического спектра предполагается регистрировать в модельном или реальном эксперименте. Например, степенной спектр, или несколько монохроматических пучков фиксированной энергии. Соответственно выбрать форму обучающей и контрольной выборок для набора последующего набора компьютерной статистики. Кроме того, выбрать тип погрешности, которая будет характеризовать точность методики энергетических измерений и которую мы будем минимизировать – абсолютную (3) или относительную (4);
2. Смоделировать на входе компьютерной модели детектора пучок первичных частиц интересующего нас сорта выбранной ранее формы энергетического спектра. Чем больше будет размер этого «обучающего» пучка, тем точнее будут оценены все необходимые для определения энергии корреляционные связи, но на практике можно ограничиться статистикой, объем  $n$  которой как минимум имеет тот же порядок, что и количество измеряемых детектором параметров, или превышает его хотя бы в несколько раз. Для спектрометра KLEM, в котором используются две перпендикулярно ориентированные плоскости шириной 6 см каждая, количество измеряемых параметров при толщине каждого стрипа 50 мкм равно  $m=2*6*10^{-2}/(50*10^{-6})=2400$  и при толщине стрипа 150 мкм  $m=800$ . Записать в файл последовательно для

каждого из смоделированных актов прохождения частиц пучка через прибор заданные значения первичной энергии  $E$  и измеренные значения ионизационных потерь  $N_i$ . Сформировать обучающую выборку  $\{E, \{N_i\}\}_1, \{E, \{N_i\}\}_2, \dots, \{E, \{N_i\}\}_n$  и преобразовать все  $N_i$  в  $N_i^a$ , где константа  $a$  пока неизвестна;

3. Определить значения неизвестных констант в формуле (2) оценивания первичной энергии (в случае минимизации абсолютной погрешности методика описана, например, в [2], а нацеленный на относительную погрешность алгоритм представляет собой ее модификацию и сводится к ней заменой переменных);
4. Сформировать тестирующую выборку аналогично пункту 1, для каждой первичной частицы из этой выборки оценить энергию и рассчитать для нескольких значений  $a$  усредненную по всей выборке погрешность (3) или (4). Определить, при каком значении  $a$  погрешность минимальна, запомнить это значение и использовать его в дальнейшем для повышения точности методики за счет учета нелинейности статистической связи между энергией и ионизационными потерями. В наших вычислительных экспериментах оптимальное значение  $a$  составило  $a \approx 1,25-1,3$ ; впрочем, не исключено, что оно зависит от толщины стрипа и от других особенностей установки.
5. Формула оценивания первичной энергии (2) уже полностью определена, поскольку в предыдущих пунктах найдены значения всех входящих в нее констант. Остается только подать на вход компьютерной модели измерительной аппаратуры (а в дальнейшем – и реального прибора) контрольный пучок интересующей нас структуры, спектральной формы и интенсивности, оценить для каждой частицы пучка первичную энергию и посчитать погрешность ее определения, тем самым проверив качество созданной методики.

### Результаты компьютерных экспериментов

Приведем результаты применения описанной выше методики к решению двух задач в рамках одной общей задачи восстановления энергии первичных частиц многомерным методом, которой посвящена первая часть работы. Первая из них – определение энергии каждой частицы из пучка степенного энергетического спектра, вторая – определение энергии частиц из нескольких монохроматических по энергии пучков от  $E = 10^{11}$  до  $E = 10^{16}$  эВ.

Во всех рассмотренных случаях моделировались только протоны, падающие на плоскость измерительной аппаратуры (плоскость мишени, см. рис.1) под прямым углом. Остальные параметры модельных экспериментов: размеры микрострипового детектора  $6 \times 6$  см, толщина мишени из углерода 10 см, высота установки 31 см, толщина конвертора 2 см. Для моделирования использовался программный пакет GEANT 3.21 [6]. При этом высокоэнергетичные адронные взаимодействия описывались с помощью генератора QGSJET [7], низкоэнергетичные (до 50 ГэВ) – с помощью генератора FLUKA [6]. Применимость моделей для описания адронных взаимодействий подтверждена сравнением с экспериментальными данными [7, 8].

В первом случае предполагается, что нам известна форма реального энергетического спектра – степенная функция, хотя, быть может, и с неточно известным показателем степени. Для того, чтобы учесть эту известную нам априорную информацию, следует использовать для обучения нашей методики случайную выборку именно из степенного распределения с предполагаемым значением степенного показателя  $\gamma$ . В качестве критерия была выбрана относительная, а не абсолютная погрешность. Результаты моделирования таковы: с.к. погрешность составляет 49% для толщины стрипов детектора и в 50, и в 150 мкм (разница составляет полпроцента). Одномерный метод, использующий параметр  $S$ , дает 54% и 56% соответственно. Однако, повторим, в этих цифрах подавляющий вклад вносят события малой энергии, поскольку они являются результатом усреднения по круто падающему степенному спектру.

Во втором случае, когда контрольной выборкой являются монохроматические пучки от  $E = 10^{11}$  до  $E = 10^{16}$  эВ, нас «в равной степени» интересуют значения энергии во всем диапазоне по логарифмической шкале от  $10^{11}$  до  $E = 10^{16}$ , а не только те шесть «круглых» значений, которые уложились в этот интервал. Это означает, что обучающей следует сделать случайную выборку из такого энергетического спектра, в котором логарифм энергии распределен равномерно на всем интересующем интервале. Подобная обучающая выборка была наиграна объемом всего около 500 событий на всем энергетическом интервале, охватывающем пять порядков. Объем каждой из шести монохроматических по энергии контрольных выборок колебался от 100 до 500 событий для  $10^{11} - 10^{15}$  эВ и около 40 событий – для  $10^{16}$ .

Результаты оценивания энергии приведены в таблице 1 для нескольких значений описывающего нелинейность задачи параметра  $a$ . Для сравнения с уже существующей и использующейся при моделировании энергетических измерений в проекте KLEM-NUCLEON методикой оценивания энергии аналогичные погрешности были вычислены на тех же контрольных выборках с помощью алгоритма по одному параметру  $S$  (1). Они приведены в предпоследней строке таблицы. Отметим, что при вычислении всех этих погрешностей, в отличие от того, как это делалось ранее, не было проведено никакого «усечения» по низким энергиям, при усреднении была использована вся насчитанная для контрольных выборок статистика. Кроме того, существенно, что погрешности прежней методики рассчитаны фактически на тех же данных, которые были использованы в самом алгоритме для построения калибровочной кривой (таким образом, «восстанавливалась» уже априори известная энергия). Это означает, что на других выборках следует ожидать меньшую точность восстановления энергии в рамках прежнего алгоритма, особенно при энергии  $10^{16}$ , так как соответствующая статистика очень мала (около 40 событий), что связано с большим временем обработки данных.

В самой нижней строке мы приводим ранее публиковавшиеся [5] результаты компьютерных экспериментов по оцениванию первичной энергии спектрометром KLEM, когда погрешность вычислялась не по всей статистике, а с выброшенной из нее наиболее низкоэнергетической частью. Будучи некорректной с математической

точки зрения, эта процедура тем не менее не приводила к потере информации, существенной для дальнейшей обработки данных, поскольку далее восстановленные значения энергии напрямую использовались для построения первичного спектра с помощью гистограмм. Область низких энергий не представляет интереса с этой точки зрения, тогда как «хвост» в области заниженных значений  $E_{\text{оцен.}}$  вносит значительный вклад в общую погрешность, что можно заметить из сравнения двух последних строк таблицы. В данной же работе будет предложен алгоритм восстановления показателя степенного спектра, не использующий значения оценок энергии, а только оценку  $\langle \ln E \rangle$ .

Как видно из таблицы, предложенный метод дает значительное уменьшение ошибки в определении энергии. Отметим также, что уменьшение разрешающей способности стриповых детекторов с 50 до 150 мкм (при неизменных размерах всего прибора, что дает возможность сократить количество стрипов в 3 раза) увеличивает погрешность весьма незначительно.

**Табл. 1.** Относительная погрешность восстановления энергии для двух вариантов разрешающей пространственной способности стриповых детекторов  $dx=50$  мкм и  $dx=150$  мкм.

$E$ , эВ	$10^{11}$	$10^{12}$	$10^{13}$	$10^{14}$	$10^{15}$	$10^{16}$
<b>Многомерный метод</b>						
$a = 1,2$						
$dx=50$ мкм	46%	58%	61%	60%	63%	71%
$dx=150$ мкм	46%	64%	68%	61%	63%	72%
$a = 1,3$						
$dx=50$ мкм	43%	59%	61%	62%	66%	65%
$dx=150$ мкм	43%	60%	65%	63%	66%	67%
$a = 1,4$						
$dx=50$ мкм	37%	62%	62%	63%	69%	61%
$dx=150$ мкм	35%	62%	64%	64%	70%	64%
$a = 1,5$						
$dx=50$ мкм	33%	64%	64%	65%	73%	60%
$dx=150$ мкм	25%	65%	65%	65%	73%	62%
<b>Метод использования</b>						
<b>1 параметра <math>S</math>, <math>dx=50</math> мкм</b>						
Полная статистика	92%	103%	101%	95%	83%	61%
Результат из работы [5] без низкоэнергетической части статистики	72%	69%	61%	55%	56%	60%

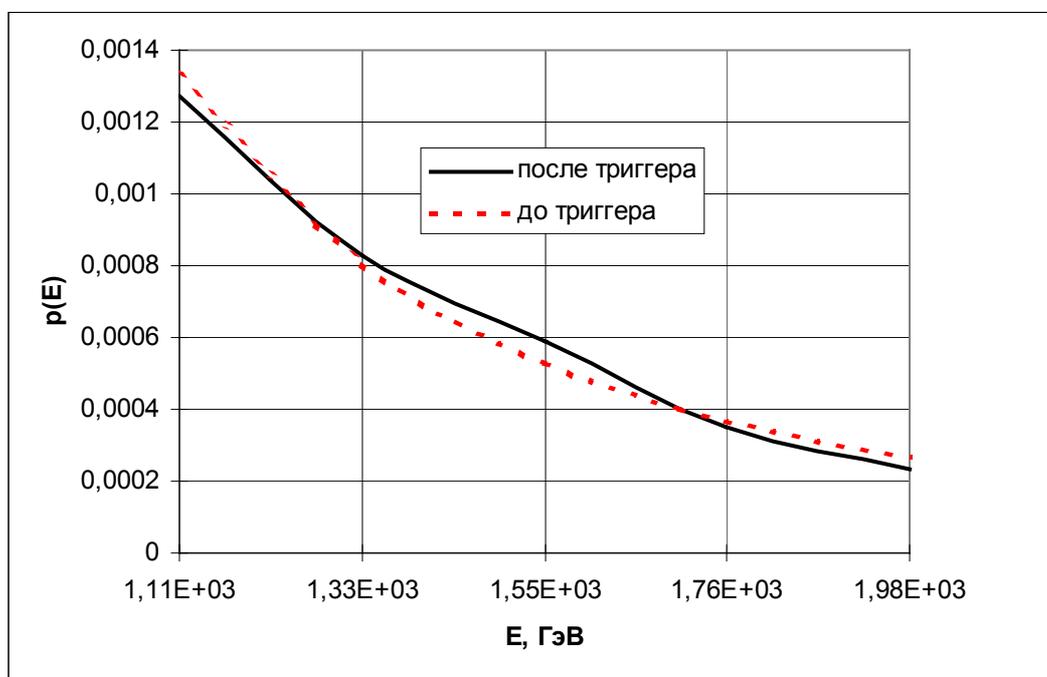
## 2. ОПРЕДЕЛЕНИЕ ПОКАЗАТЕЛЯ ПЕРВИЧНОГО СТЕПЕННОГО СПЕКТРА

В этой части работы рассматривается методика, алгоритм и результаты оценивания показателя степенного спектра первичных КЛ на основе описанного выше многомерного метода восстановления первичной энергии.

Для восстановления формы спектра нами раньше использовалась традиционная методика построения гистограмм по восстановленным значениям [9, 10]. Эта методика сопровождается большой и трудно оцениваемой погрешностью, но она давно принята на вооружение и является неким стандартом. Тем не менее, предложенный в первой части метод восстановления энергии является несмещенным и сопровождается минимальной с.к. погрешностью, поэтому вполне естественно ожидать, что даже простое применение «в лоб» этой методики к значениям энергии, восстановленным новым, а не используемым сейчас, методом, приведет к большей точности в восстановлении формы спектра.

Одна из главных сложностей восстановления спектра состоит в том, что при моделировании работы прибора KLEM производится «многоступенчатый» отбор событий, которые будут зафиксированы детекторами. В результате этого отбора форма первичного спектра очень сильно искажается, так что даже абсолютно точное измерение значений энергии зафиксированных прибором частиц не позволяет восстановить вид их спектра на входе измерительной аппаратуры (рис. 2).

**Рис. 2.** Плотность первичного распределения до триггера (аналитический вид) и после (построена на основе гистограммы)



Рассматриваемое вполне реальным условие отбора приводит к тому, что рассчитанный методом максимального правдоподобия показатель степенного первичного спектра с нижней границей  $E_0 = 1$  ТэВ занижается:

$$\gamma_0 = 2,70 \Rightarrow \gamma_{\text{оцен.}} = 2,58,$$

$$\gamma_0 = 3,00 \Rightarrow \gamma_{\text{оцен.}} = 2,88,$$

$$\gamma_0 = 3,30 \Rightarrow \gamma_{\text{оцен.}} = 3,17$$

(повторим, в случае точного измерения энергии всех прошедших триггерный отбор частиц), тогда как чисто статистическую погрешность оценивания  $\gamma$  можно грубо оценить, исходя из реализующей этот метод в случае точных измерений формулы (5)

$$\sigma_{\gamma_{\text{оцен.}}} \sim \sigma \{1 / \ln (E/E_0)\} \cong 0,017,$$

$n \sim 10000$  – объем статистики.

### **Оценивание показателя степенного спектра методом максимального правдоподобия**

В данной работе мы рассмотрим восстановление только числового параметра, характеризующего первичный спектр. Восстановление аналитической формы функции энергетического распределения по многомерной методике – вопрос дальнейших исследований. Отдельно мы рассмотрим случай степенного спектра, ограниченного не только слева, но и справа, что соответствует ситуации, когда нас интересует только определенный участок реального энергетического спектра первичных частиц.

#### ***1. Неограниченный справа спектр***

Для степенного распределения единственным параметром (при фиксированной левой границе) является  $\gamma$ , показатель степени дифференциального первичного спектра.

Для того чтобы оценить параметр  $\gamma$  распределения первичной энергии частиц, зафиксированных детектором, можно использовать метод максимального правдоподобия (ММП). Оценивание параметров спектра с помощью ММП – стандартная в физике космических лучей процедура. Как известно, в случае, если распределение степенное, оценка его параметра  $\gamma$  методом МП вычисляется по формуле:

$$\gamma_{\text{оцен.}} = 1 + 1 / (\langle \ln E \rangle - \ln E_0), \quad (5)$$

где  $\langle \ln E \rangle$  – среднее значение логарифма истинной первичной энергии,  $E_0$  – левая граница спектра. Мы можем его найти путем усреднения логарифмов измеренных значений энергии только в том случае, когда энергия измеряется абсолютно точно, либо погрешности определения  $\ln E$  не зависят от энергии. И то, и другое – идеализация. Во всех же прочих случаях распределение измеренной энергии представляет собой свертку первичного спектра с функцией, описывающей вносимые детектирующим прибором искажения. Если эта функция известна априори или хотя бы достаточно точно аппроксимирована по экспериментальным

или модельным данным, то в рамках общей идеологии оценивания параметра спектра методом МП следует вычислить свертку этих двух функций, а затем вывести для получившегося распределения новую формулу оценивания параметра  $\gamma$ , применив методику МП с самого начала и заново повторив все ее шаги, аналогичные тем, которые приводят к формуле (5). В нашем же случае возникает дополнительное искажение формы спектра еще на предварительном этапе отбора событий аппаратными триггерами. Возможно, по этой или по какой-либо другой причине попытки аппроксимировать распределение функции прибора, а затем применить изложенную выше идеологию не принесли положительных результатов. Тем не менее, задача оценивания показателя спектра решена в рамках несколько другого подхода. Создано и отлажено программное обеспечение для определения  $\gamma$ , получены результаты модельных экспериментов.

### **АЛГОРИТМ ОЦЕНИВАНИЯ ПОКАЗАТЕЛЯ ПЕРВИЧНОГО СТЕПЕННОГО СПЕКТРА**

Отметим, что, поскольку в формуле (5) присутствуют не сами значения восстановленной энергии, а их средний логарифм, то нецелесообразно сначала оценивать  $E$ , а потом брать логарифм получившихся значений. Вместо этого следует применить такой метод, который сразу дает наиболее точную оценку логарифма энергии. Именно таким методом является линейное наилучшее в с.к. оценивание  $\ln E$  как случайной величины. А поскольку вышеназванный метод, в отличие от «метода с параметром  $S$ », несмещенный, это гарантирует нам, что оценка среднего значения оцениваемого логарифма энергии не будет подвержена никаким систематическим занижениям или завышениям.

1) «Предварительный» этап.

Восстановление логарифма энергии методом линейного наилучшего в с.к. оценивания случайных векторов. В качестве обучающей выборки используется степенной спектр с заданным показателем, в качестве контрольной – степенной спектр, параметр  $\gamma$  которого неизвестен и подлежит оцениванию. Именно степенное распределение следует использовать для обучения метода потому, что результаты оценивания будут тем лучше, чем больше априорной информации мы тем или иным образом учтем – предположение же о степенном характере первичного спектра как раз является такой информацией.

2) Нахождение «предварительной» оценки  $\gamma$  по формуле (5).

3) Повторение процедуры 1)-2) для нескольких значений  $\gamma$  из интересующего нас интервала (в наших вычислительных экспериментах мы использовали всего три точки,  $\gamma_0 = 2,7$ ,  $\gamma_0 = 3,0$  и  $\gamma_0 = 3,3$ , что уже обеспечило достаточно хорошую точность).

Таким образом, для каждого из этих трех значений  $\gamma_0$  мы определили, чему равняются значения оценок первого, второго и третьего  $\gamma_0$  по формуле (5).

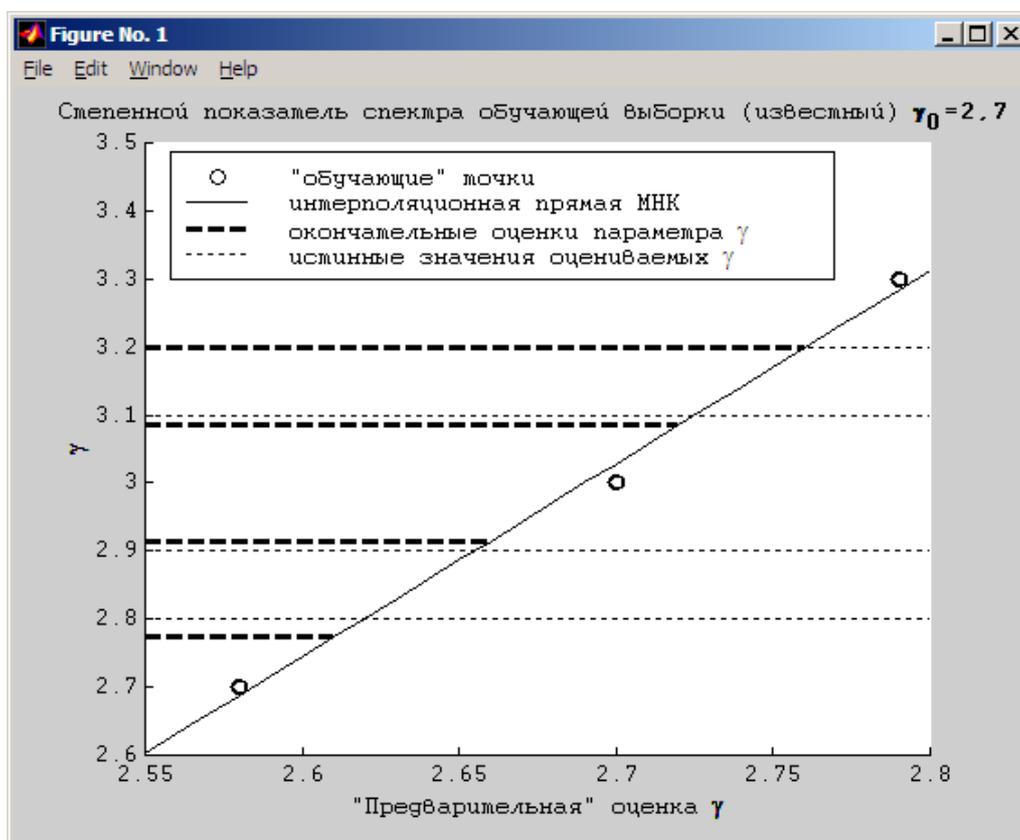
4) Построение интерполяционной кривой зависимости истинного значения показателя  $\gamma$  от его «предварительной» оценки по найденным в предыдущем пункте

точкам (в нашем случае только трех точек лучшие результаты были получены с помощью линейной функции).

Для каждого из трех использовавшихся на «предварительном этапе» значений  $\gamma_0$ , то есть для каждой из трех имеющихся пока в нашем распоряжении обучающих выборок, была построена своя интерполяционная кривая. В результате мы получили возможность сравнить качество разработанной методики для различных значений закладываемого для ее обучения  $\gamma$ , в случае, если оцениваемое значение намного больше или намного меньше известного, и т.д.

5) Нахождение оценок параметра  $\gamma$ , скорректированных с помощью каждой из трех построенных в предыдущем пункте интерполяционных кривых. Пример того, как описанная процедура изображается графически, приведен на рис. 3.

**Рис. 3.** *Оценивание  $\gamma$  по линейной зависимости  $\gamma$  ( $\gamma_{\text{предвар.}}$ )*



### Результаты компьютерных экспериментов

Были проведены компьютерные эксперименты по оцениванию параметра  $\gamma$  степенного спектра протонов (в случае вертикального падения пучка на плоскость детектора) согласно приведенному выше алгоритму. Использовались модели измерительного детектора с толщиной стрипов 50 и 150 мкм. Для сравнения многомерной и основанной на использовании одного параметра методик показатель  $\gamma$  оценивался на одних и тех же модельных данных двумя способами: по алгоритму,

использующему многомерное статистическое оценивание логарифма энергии, и методом восстановления энергии на основе «параметра  $S$ » (1).

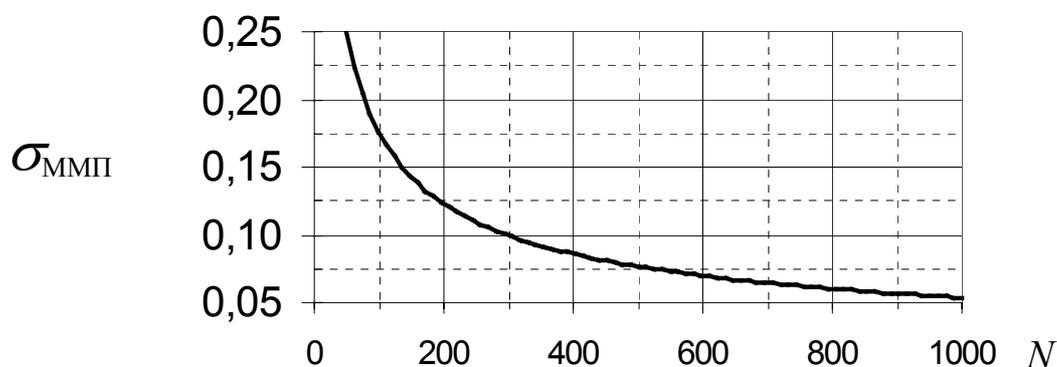
Обучение многомерного метода проводилось по трем пучкам протонов степенного энергетического спектра с показателем  $\gamma$ , равным 2,70, 3,00 и 3,30. Оценивание степенного показателя проводилось для пучков, истинное значение  $\gamma$  которых равнялось 2,8, 2,85, 2,9, 3,1, 3,15 и 3,2. Количество событий в каждом из обучающих пучков было достаточно большим (несколько тысяч), это не накладывает никаких ограничений для практической реализации тестируемой методики, поскольку обучающие выборки вполне возможно получать не в реальном эксперименте, дающем малую статистику, а компьютерным моделированием, которое позволяет наиграть спектры любой формы и объема.

Контрольные выборки были взяты объемом от 100 до 400 событий – такое количество первичных протонов вполне может быть зафиксировано установкой KLEM на борту космического аппарата.

Результаты оценивания приведены в таблице 2. Для наглядности первая графа ( $N = 100$ ) представлена графически на рис. 5.

В соответствующих ячейках таблицы можно видеть, насколько сильно влияет на точность оценивания объем набранной статистики. Как было отмечено выше, статистическая погрешность оценки, обусловленная *исключительно конечностью выборки* (энергия известна точно), оценивается из формулы метода МП (5). Она является как бы «неустранимой», поэтому очень важно иметь представление о порядке ее величины, как о «предельной» (соответствующей случаю идеально точных измерений) «разрешающей способности» методики (или прибора) восстановления параметра  $\gamma$ . График зависимости данной «неустранимой» погрешности от объема статистики  $N$  приведен на рис. 4 для интересующего нас интервала в районе нескольких сот событий. Ее значения, соответствующие каждому из  $N$  – в левом столбце таблицы.

**Рис. 4.** Зависимость погрешности ММП оценивания  $\gamma$  при точно известной первичной энергии от объема статистики



Отметим следующее. Во-первых, хотя в каждом «акте» оценивания  $\gamma$  по формуле (5) задействованы данные только  $N$  (от 100 до 400) событий, но каждый из табличных результатов усреднен по нескольким  $N$ -событийным выборкам среди

всей имеющейся статистики, это сделано для того, чтобы абстрагироваться от случайных «выбросов» и проверить несмещенность методики, отсутствие систематических сдвигов. Однако в реальном эксперименте у нас будет небольшая статистика (как раз порядка нескольких сотен событий), что не даст нам возможности усреднить результаты, полученные по нескольким выборкам. Поэтому не менее ценную информацию несут значения погрешности в определении  $\gamma$  (они даже выделены в таблице жирным шрифтом), которые говорят нам о том, насколько большие отклонения (от приведенного слева от погрешности усредненного значения) можно ожидать при проведении однократного эксперимента на статистике в  $N$  событий. Качество методики оценивания тем лучше, в чем меньшее количество раз эта погрешность превышает «неустранимую» погрешность, содержащуюся в методе МП и представленную в левом столбце таблицы 2 под статистикой и на рис. 4.

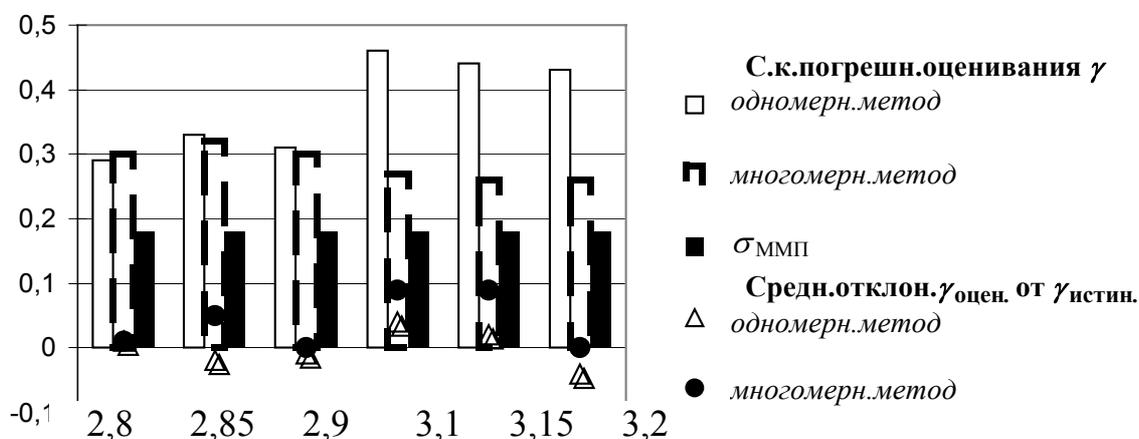
Табл. 2. Значения оценок параметра  $\gamma$

		Истинные значения $\gamma$ →		2,8		2,85		2,9		3,1		3,15		3,2		
Статистика		$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$	
$N = 100,$ $\sigma_{\text{ММП}} = 0,18$	«Метод с парам.S»	2,81	<b>0,29</b>	2,90	<b>0,33</b>	2,90	<b>0,31</b>	3,19	<b>0,46</b>	3,24	<b>0,44</b>	3,20	<b>0,43</b>			
	Новый	$\gamma_0 = 2,7$	2,81	<b>0,30</b>	2,83	<b>0,32</b>	2,89	<b>0,30</b>	3,14	<b>0,27</b>	3,17	<b>0,26</b>	3,16	<b>0,26</b>		
		$\gamma_0 = 3,0$	2,81	<b>0,31</b>	2,85	<b>0,31</b>	2,91	<b>0,29</b>	3,16	<b>0,26</b>	3,17	<b>0,25</b>	3,16	<b>0,25</b>		
		$\gamma_0 = 3,3$	2,81	<b>0,30</b>	2,85	<b>0,30</b>	2,91	<b>0,30</b>	3,13	<b>0,27</b>	3,17	<b>0,24</b>	3,17	<b>0,26</b>		
$N = 200,$ $\sigma_{\text{ММП}} = 0,12$	«Метод с парам.S»	2,80	<b>0,21</b>	2,88	<b>0,22</b>	2,92	<b>0,24</b>	3,11	<b>0,29</b>	3,22	<b>0,34</b>	3,13	<b>0,25</b>			
	Новый	$\gamma_0 = 2,7$	2,80	<b>0,22</b>	2,84	<b>0,22</b>	2,89	<b>0,21</b>	3,14	<b>0,19</b>	3,18	<b>0,18</b>	3,15	<b>0,18</b>		
		$\gamma_0 = 3,0$	2,80	<b>0,22</b>	2,84	<b>0,22</b>	2,90	<b>0,21</b>	3,14	<b>0,19</b>	3,17	<b>0,17</b>	3,15	<b>0,19</b>		
		$\gamma_0 = 3,3$	2,80	<b>0,22</b>	2,83	<b>0,22</b>	2,89	<b>0,21</b>	3,14	<b>0,19</b>	3,17	<b>0,18</b>	3,14	<b>0,19</b>		
$N = 300,$ $\sigma_{\text{ММП}} = 0,10$	«Метод с парам.S»	2,81	<b>0,15</b>	2,84	<b>0,13</b>	2,88	<b>0,16</b>	3,12	<b>0,20</b>	3,20	<b>0,25</b>	3,09	<b>0,21</b>			
	Новый	$\gamma_0 = 2,7$	2,79	<b>0,17</b>	2,83	<b>0,18</b>	2,89	<b>0,17</b>	3,13	<b>0,15</b>	3,17	<b>0,15</b>	3,16	<b>0,15</b>		
		$\gamma_0 = 3,0$	2,80	<b>0,18</b>	2,83	<b>0,18</b>	2,89	<b>0,17</b>	3,13	<b>0,15</b>	3,17	<b>0,15</b>	3,14	<b>0,15</b>		
		$\gamma_0 = 3,3$	2,80	<b>0,18</b>	2,84	<b>0,18</b>	2,89	<b>0,17</b>	3,13	<b>0,16</b>	3,17	<b>0,15</b>	3,15	<b>0,15</b>		
$N = 400,$ $\sigma_{\text{ММП}} = 0,09$	«Метод с парам.S»	2,78	<b>0,13</b>	2,86	<b>0,13</b>	2,87	<b>0,14</b>	3,15	<b>0,19</b>	3,16	<b>0,19</b>	3,13	<b>0,19</b>			
	Новый	$\gamma_0 = 2,7$	2,79	<b>0,15</b>	2,84	<b>0,16</b>	2,89	<b>0,16</b>	3,13	<b>0,13</b>	3,18	<b>0,13</b>	3,15	<b>0,14</b>		
		$\gamma_0 = 3,0$	2,80	<b>0,16</b>	2,83	<b>0,16</b>	2,89	<b>0,14</b>	3,12	<b>0,14</b>	3,18	<b>0,13</b>	3,15	<b>0,14</b>		
		$\gamma_0 = 3,3$	2,80	<b>0,16</b>	2,74	<b>0,16</b>	2,89	<b>0,15</b>	3,12	<b>0,14</b>	3,17	<b>0,13</b>	3,15	<b>0,14</b>		

Отметим и то обстоятельство, что, как и в задаче оценивания энергии индивидуальных частиц, толщина стрипов детектора оказывает очень незначительное влияние на величину погрешности многомерной методики. В данном случае как сами оценки показателя, так и их дисперсии настолько слабо (порядка 0,01) изменялись при увеличении толщины стрипа в три раза, с 50 до 150 мкм (количество стрипов во всей модели детектора при этом уменьшается в три раза и соответственно появляется возможность снижения затрат на изготовление аппаратуры), что во всех таблицах приводятся только результаты для 150 мкм.

Насчет сравнения двух методик, многомерной и однопараметрической, отметим, что оно здесь проведено весьма условно, поскольку никакого определенного алгоритма восстановления показателя степенного спектра не существует, что, кстати, и является одной из причин для разработки нового универсального алгоритма, применимого, как и лежащий в его основе многомерный метод оценивания первичной энергии, для любого многопараметрического детектора. В нашем случае мы просто подвергли абсолютно одинаковой обработке значения  $E$  (или  $\ln E$ ) каждой из первичных частиц промоделированных пучков, полученные как многомерным методом, так и одномерным «с параметром  $S$ ». Поэтому фактически к результатам энергетических измерений обоими методиками затем применялся один и тот же комплексный алгоритм обработки.

**Рис. 5.** Отклонения от среднего и с.к. погрешности оценок параметра  $\gamma$ .  
( $N = 100$ ,  $\sigma_{\text{ММП}} = 0,18$ )



### Оценивание $\gamma$ при неизвестной левой границе спектра

Для того, чтобы приблизить условия вычислительного эксперимента к реальности, когда нам не будет точно известна левая граница спектра регистрируемых частиц, мы отдельно промоделировали работу нашего метода для этого случая. В этой ситуации значение  $E_0$  оценивалось, т.е. работа происходила только с событиями, отобранными согласно критерию  $E_{\text{изм.}} > E_0$ , где  $E_0$  – известная задаваемая величина. (Говоря более строго, в рамках предлагаемой многомерной

методики оценивания сама энергия не определяется, поэтому условие отбора в данном алгоритме выглядело как  $\ln E_{\text{изм.}} > \ln E_0$ ). Значение левой границы спектра выбиралось из соображений, во-первых, возможности его достаточно точного оценивания, и во-вторых, размера необходимой для работы статистики, поскольку именно от объема статистики зависит величина «неустранимой» статистической погрешности оценивания  $\gamma$ . Из этих соображений было выбрано несколько значений  $E_0$  в диапазоне от 2 ТэВ до 4 ТэВ. Результаты оценивания  $\gamma$  для некоторых  $E_0$  приведены в таблице 3.

**Табл. 3.** Значения оценок параметра  $\gamma$  (при оценивании левой границы спектра,  $N = 100, \sigma_{\text{ММП}} = 0,18$ )

Истинные значения $\gamma \rightarrow$		2,8		2,85		2,9		3,1		3,15		3,2	
$E_0$	$\gamma_0$	$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$										
2 ТэВ	2,7	2,84	<b>0,23</b>	2,83	<b>0,23</b>	2,91	<b>0,24</b>	3,11	<b>0,23</b>	3,25	<b>0,28</b>	3,19	<b>0,24</b>
	3,0	2,82	<b>0,23</b>	2,83	<b>0,24</b>	2,91	<b>0,23</b>	3,10	<b>0,24</b>	3,24	<b>0,26</b>	3,17	<b>0,23</b>
	3,3	2,83	<b>0,22</b>	2,80	<b>0,23</b>	2,88	<b>0,22</b>	3,08	<b>0,24</b>	3,22	<b>0,25</b>	3,16	<b>0,23</b>
2,5 ТэВ	2,7	2,85	<b>0,25</b>	2,81	<b>0,24</b>	2,90	<b>0,24</b>	3,14	<b>0,25</b>	3,27	<b>0,30</b>	3,26	<b>0,26</b>
	3,0	2,86	<b>0,24</b>	2,79	<b>0,23</b>	2,91	<b>0,24</b>	3,10	<b>0,23</b>	3,26	<b>0,29</b>	3,30	<b>0,27</b>
	3,3	2,86	<b>0,27</b>	2,76	<b>0,25</b>	2,92	<b>0,25</b>	3,11	<b>0,24</b>	3,33	<b>0,33</b>	3,30	<b>0,28</b>
3 ТэВ	2,7	2,83	<b>0,22</b>	2,80	<b>0,22</b>	2,93	<b>0,23</b>	2,97	<b>0,22</b>	3,26	<b>0,27</b>	3,26	<b>0,23</b>
	3,0	2,83	<b>0,19</b>	2,77	<b>0,20</b>	2,93	<b>0,21</b>	2,97	<b>0,20</b>	3,23	<b>0,23</b>	3,20	<b>0,19</b>
	3,3	2,80	<b>0,21</b>	2,80	<b>0,21</b>	2,93	<b>0,21</b>	3,02	<b>0,18</b>	3,27	<b>0,27</b>	3,20	<b>0,19</b>

Сравнивая с данными в таблице 2, видим, что показатель спектра, левая граница которого априори неизвестна и восстанавливается по результатам измерений регистрирующей аппаратуры, оценивается с не меньшей точностью, чем спектра с фиксированным  $E_0$ .

## 2. Спектр на участке ( $E_1, E_2$ )

Если нас интересует значение показателя  $\gamma$  только на определенном участке всего неограниченного справа энергетического спектра первичных КЛ, мы тем самым переходим к задаче оценивания параметра  $\gamma$  иного, чем в решенной выше задаче, спектра, а именно

$$p(E) = \begin{cases} C(\gamma)\exp(-\gamma), & E_1 < E < E_2, \\ 0, & E < E_1, E > E_2 \end{cases}$$

Эта форма спектра немного более сложна с точки зрения оценивания  $\gamma$ , потому что метод максимального правдоподобия в данном случае дает не готовую формулу (5), а нелинейное уравнение относительно  $\gamma$ , корнем которого и является  $\gamma_{\text{оцен.}}$ :

$$\gamma = 1 + (E_1^{1-\gamma} - E_2^{1-\gamma}) / (E_1^{1-\gamma} \{ \langle \ln E \rangle_N - \ln E_1 \} - E_2^{1-\gamma} \{ \langle \ln E \rangle_N - \ln E_2 \} ). \quad (6)$$

## АЛГОРИТМ ОЦЕНИВАНИЯ ПОКАЗАТЕЛЯ ПЕРВИЧНОГО СТЕПЕННОГО СПЕКТРА НА ОГРАНИЧЕННОМ УЧАСТКЕ

Алгоритм почти в точности повторяет описанный выше, для случая неограниченного спектра, за тем исключением, что лежащий в их основе метод МП теперь реализуется не формулой (5), а уравнением (6).

1) «Предварительный» этап.

Восстановление логарифма энергии методом линейного наилучшего в с.к. оценивания случайных векторов. В качестве обучающей выборки используется степенной спектр с заданным показателем, в качестве контрольной – степенной спектр, параметр  $\gamma$  которого неизвестен и подлежит оцениванию.

2) Нахождение численным методом «предварительной» оценки  $\gamma$  из нелинейного уравнения (6).

3) Повторение процедуры 1)-2) для нескольких значений  $\gamma$  из интересующего нас интервала (в наших вычислительных экспериментах мы использовали всего три точки,  $\gamma_0 = 2,7$ ,  $\gamma_0 = 3,0$  и  $\gamma_0 = 3,3$ ).

4) Построение интерполяционной кривой зависимости истинного значения показателя  $\gamma$  от его «предварительной» оценки по найденным в предыдущем пункте точкам.

5) Нахождение оценок параметра  $\gamma$ , скорректированных с помощью каждой из трех построенных в предыдущем пункте интерполяционных кривых.

## Результаты компьютерных экспериментов

Были проведены компьютерные эксперименты по оцениванию параметра  $\gamma$  степенного спектра протонов на различных энергетических интервалах. Как и для неограниченного спектра, рассматривался случай вертикального падения пучка на плоскость детектора, использовалась модель измерительного детектора с толщиной стрипов 150 мкм. Также было проведено сравнение многомерной методики с традиционной, использующей для оценивания первичной энергии индивидуальных событий один параметр  $S$  (1). Все прочие параметры моделирования точно такие же, как в предыдущем случае. Значения как левой, так и правой границы спектра не считались известными и восстанавливались по модельным данным, как в табл. 3 предыдущего раздела.

Даже для такой маленькой статистики, как 100 событий, хорошо видно, что оценивание  $\gamma$  на энергетическом интервале небольшой ширины предлагаемой методикой вполне допустимо, хотя и сопровождается чуть большей, чем для неограниченного спектра, погрешностью. (Напомним, границы интервала берутся по восстановленной первичной энергии, потому что ограничения по истинной, неизвестной нам энергии не могут быть соблюдены в реальном физическом

эксперименте.) Это объясняется тем, что бóльшую величину имеет и «неустраняемая» погрешность, «сокрытая» в методе максимального правдоподобия, который теперь имеет более сложный алгоритм – нелинейное уравнение (6) вместо простой готовой формулы (5) для неограниченного спектра. Теперь величина этой погрешности зависит не только от объема статистики, но и от ширины интервала. Для представленного в таблице 4 случая (интервал в 2 ТэВ) «неустраняемая» погрешность ММП составляет 0,29, так что соотношение фактической и минимально возможной погрешностей остается примерно таким же, как и при оценивании  $\gamma$  неограниченного спектра.

**Табл. 4.** Значения оценок параметра  $\gamma$  на интервале  $2 \text{ ТэВ} < E < 4 \text{ ТэВ}$ .

$$N = 100, \sigma_{\text{ММП}} = 0,29$$

Истинные значения $\gamma$ →		2,8		2,85		2,9		3,1		3,15		3,2	
		$\hat{\gamma}_{\text{оцен.}}$	$\Delta\gamma$										
«Метод с парам. S»		2,94	<b>0,50</b>	2,93	<b>0,45</b>	2,86	<b>0,46</b>	2,94	<b>0,48</b>	3,10	<b>0,48</b>	2,95	<b>0,52</b>
Новый	$\gamma_0 = 2,7$	2,78	<b>0,37</b>	2,85	<b>0,39</b>	2,82	<b>0,41</b>	3,18	<b>0,40</b>	3,19	<b>0,41</b>	3,18	<b>0,40</b>
	$\gamma_0 = 3,0$	2,72	<b>0,36</b>	2,94	<b>0,37</b>	2,87	<b>0,37</b>	3,14	<b>0,36</b>	3,15	<b>0,36</b>	3,11	<b>0,37</b>
	$\gamma_0 = 3,3$	2,70	<b>0,35</b>	2,77	<b>0,36</b>	2,74	<b>0,39</b>	3,04	<b>0,37</b>	3,07	<b>0,38</b>	3,00	<b>0,41</b>

## Заключение

Основные полученные в работе результаты можно сформулировать следующим образом.

Разработаны, реализованы, протестированы и исследованы в компьютерных экспериментах *новые методики* обработки результатов измерений детекторов, позволяющих регистрировать значения большого количества параметров:

- методика оценивания недоступных непосредственному измерению физических параметров (таких, как первичная энергия и другие характеристики первичной частицы);
- методика оценивания показателя первичного степенного спектра КЛ (для «неограниченного» спектра на полуинтервале и для участка спектра от  $E_1$  до  $E_2$ ).

Данные методики представляют собой оптимальные в рамках широкого класса алгоритмов процедуры, т.е. сопровождаются наивысшей среди этого класса методов точностью.

Новые оценивающие методики исследованы в компьютерных экспериментах с математической моделью измерительной аппаратуры KLEM из проекта NUCLEON. Показано, что:

- при оценивании энергии первичных частиц новая методика дает значительно (в полтора раза) меньшую погрешность по сравнению с использовавшимся нами ранее одномерным алгоритмом.

Однако необходимо отметить, что сам по себе этот факт не приводит к такому же уменьшению погрешности восстановления спектра, поскольку представленный в работе алгоритм оценивания спектрального степенного показателя не использует восстановленные значения первичной энергии, в процессе оценивания непосредственно определяя значение  $\langle \ln E \rangle$ ;

- при оценивании показателя спектра первичных КЛ предлагаемая методика работает как минимум не хуже аналогичной одномерной, обладая при этом, в отличие от использовавшихся нами ранее в проекте KLEM-NUCLEON и других исследованиях алгоритмов определения показателя по гистограммам на основе одномерного оценивания энергии, строгой математической оптимальностью и универсальностью и позволяя исследователю применять однотипную математическую программу для обработки многопараметрических данных любого рода.

В ближайшем будущем в рамках изложенной в данной работе идеологии будет реализован алгоритм выявления неоднородностей спектра первичных КЛ. Оптимальное решение означенной задачи – предмет отдельного исследования, поскольку оно вовсе необязательно должно базироваться на решении совершенно другой с математической точки зрения задачи восстановления числового параметра – степенного показателя спектра (точно так же, как описанное в нашей работе оптимальное решение задачи определения спектрального показателя не использует в качестве входных данных восстановленные значения энергии индивидуальных событий). Речь будет идти о функциональном виде спектров и неоднородностей, поэтому, скорее всего, для решения будет привлечена теория проверки статистических гипотез.

## Литература

1. Адамс Дж., Башинджагян Г.Л., Башинджагян П.Г., Голубков С.А., Друри Л., Егоров Н.Н., Короткова Н.А., Подорожный Д.М., Прокурор Ж., Роганова Т.М., Сааведра О., Саймон М., Свешникова Л.Г., Сидоров А.И., Томпсон А., Турундаевский А.Н., Чилингарян А.А., Яшин И.В. Изв.АН. Сер.физ. 2001, т.65, №3. С.430-432.
2. Пытьев Ю.П. Методы анализа и интерпретации результатов эксперимента. Москва. Изд. МГУ, 1990.

3. Постников Е.Б., Башинджагян Г.Л., Короткова Н.А., Подорожный Д.М., Роганова Т.М., Свешникова Л.Г., Турундаевский А.Н. Изв.АН. Сер.физ. 2002, т.66, №11, С.1634-1637.
4. Башинджагян Г.Л., Короткова Н.А., Панасюк М.И., Подорожный Д.М., Роганова Т.М., Свешникова Л.Г., Турундаевский А.Н., Яшин И.В. Препринт № 99 - 13/571. НИИЯФ МГУ. 1999.
5. Короткова Н.А., Подорожный Д.М., Постников Е.Б., Роганова Т.М., Свешникова Л.Г., Турундаевский А.Н. Ядерная физика. 2002. Т.65. №5. С. 884-892.
6. GEANT User's Guide. CERN DD/EE/83/1. Geneva, 1983.
7. Калмыков Н.Н., Остапченко С.С. Препринт № 98-36/537. НИИЯФ МГУ. 1998
8. Рапопорт И.Д., Турундаевский А.Н., Шестоперов В.Я. Ядерная физика. 2002. Т.65. №1. С. 176.
9. Aranasenko A.V., Sukhadolskaya V.A., Derbina V.A. et al. 2001 Astrop. Phys., vol. 16, p. 13
10. Ivanenko, I.P., Shestoperov, V.Ya. et al. Proc. 23<sup>rd</sup> ICRC, 1993, vol. 2, p. 17.

## Содержание

Введение.....	3
Измерительная аппаратура KLEM.....	3
Метод линейного наилучшего в среднеквадратичном (с.к.) оценивания случайных векторов.....	4
1. ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ ЭНЕРГИИ В КАЖДОМ ИНДИВИДУАЛЬНОМ СОБЫТИИ.....	6
Алгоритм оценивания первичной энергии по данным измерительной аппаратуры.....	8
Результаты компьютерных экспериментов.....	9
2. ОПРЕДЕЛЕНИЕ ПОКАЗАТЕЛЯ ПЕРВИЧНОГО СТЕПЕННОГО СПЕКТРА.....	12
Оценивание показателя степенного спектра методом максимального правдоподобия.....	13
1. Неограниченный справа спектр.....	13
Алгоритм оценивания показателя первичного степенного спектра.....	14
Результаты компьютерных экспериментов.....	15
Оценивание $\gamma$ при неизвестной левой границе спектра.....	18
2. Спектр на участке $(E_1, E_2)$ .....	19
Алгоритм оценивания показателя первичного степенного спектра на ограниченном участке.....	20
Результаты компьютерных экспериментов.....	20
Заключение.....	21
Литература.....	22

**Дмитрий Михайлович Подорожный, Евгений Борисович Постников,  
Любовь Георгиевна Свешникова, Андрей Николаевич Турундаевский**

**Применение многомерных статистических методов  
для интерпретации данных измерений многоканальной аппаратуры  
на примере спектрометра KLEM**

Препринт НИИЯФ МГУ - 2003-12/725

Работа поступила в ОНТИ 10.06. 2003 г.

**ИД № 00545 от 06.12.1999**

**Издательство УНЦ ДО**

117246, Москва, ул. Обручева, 55А  
119992, Москва, Ленинские горы, ГЗ МГУ, Ж-105а  
Тел./факс (095) 718-6966, 939-3934  
e-mail: [izdat@abiturcenter.ru](mailto:izdat@abiturcenter.ru)  
<http://www.abiturcenter.ru/izdat/>

Гигиенический сертификат №77.99.02.923.Д.001743.03.03 от 11.03.2003  
Налоговые льготы - Общероссийский классификатор продукции  
ОК-005-93, том 1 - 953000

Заказное. Подписано в печать 10.06.2003г. Формат 60x90/16  
Бумага офсетная № 2. Усл.печ.л. 1,5  
Тираж 25 экз. Заказ № 380

Отпечатано в Мини-типографии УНЦ ДО  
в полном соответствии с качеством  
предоставленного оригинал-макета