

2001/00112

Препринт ЕФИ-940(91)-86

ԵՐԵՎԱՆԻ ՖԻԶԻԿԱՅԻ ԻՆՍՏԻՏՈՒՏ
ЕРЕВАНСКИЙ ФИЗИЧЕСКИЙ ИНСТИТУТ

С.Х. АРУТЮНЯН, Р.П. КОКОУЛИН, А.А. ЧИЛИНГАРЯН

**О НЕПАРАМЕТРИЧЕСКИХ МЕТОДАХ ПОДАВЛЕНИЯ
ФОНА ЯДЕРНЫХ КАСКАДОВ ПРИ ИЗМЕРЕНИИ СПЕКТРА
МЮОНОВ С ПОМОЩЬЮ ПАРАМЕТРА**

ЦНИИАтоминформ

ЕРЕВАН-1986

Ս.Բ. ՀԱՐՈՒԹՅՈՒՆՅԱՆ, Բ.Պ. ԿՈԿՈՌԻԼԻՆ, Ա.Ա. ԶԻԼԻՆԳԱՐՅԱՆ

ՊԱՐԵՆՏՐՈՒՄ ՄՅՈՒՈՆՆԵՐԻ ՍԳԵԿՏՐԻ ԶԱՌՈՒՄՆԵՐՈՒՄ ՄԻՋՈՒԿԵՑԻՆ
ԱՆՆԳՎԱԹԱԹԻ ՀԻՄՔԻ ԸՆՇՄԱՆ ԱԶՊԱՐԱՄՆԵՏՐԻԱՎԱՆ ԵՎԱՆԱԿՆԵՐԻ
ՄԱՍԻՆ

Մյուռնների էներգետիկական սպեկտրը պարմետր տիպի բազմաշարք, գրանցող սարքով որոշելու խնդրում առաջարկվում է նոր, ոչպարամետրիկական եղանակ մյուռնների միջուկային փոխազդեցություն ղեպքերի նույնացման համար: Առաջարկված եղանակը օգտագործում է կույտային /քլաստերային/ վերլուծության մոտեցումը՝ բնորոշող կառուցվածքները ի հայտ բերելը և հավակնում է անցման վեկտորի բաղադրիչների հաջորդականությունում առկա տարբերիչ տեղեկության, այն է՝

- ա/ էներգասնջատումը յուրաքանչյուր շարքում,
- բ/ էներգասնջատման բաշխման տեսքը ըստ սարքի երկարության, առավել լիավին ընդգրկմանը:

Երևանի ֆիզիկայի ինստիտուտ

ԵՐԵՎԱՆ 1986

УДК 537.591:518

С.Х. АРУТЮНЯН, Р.П. КОЖОУЛИН, А.А. ЧИЛИНГАРЯН

**О НЕПАРАМЕТРИЧЕСКИХ МЕТОДАХ ПОДАВЛЕНИЯ
ФОНА ЯДЕРНЫХ КАСКАДОВ ПРИ ИЗМЕРЕНИИ СПЕКТРА
МЮОНОВ С ПОМОЩЬЮ ПАРМЕТРА**

Предложен новый непараметрический метод идентификации событий ядерного взаимодействия мюонов в задаче измерения спектра мюонов многорядной детектирующей установкой типа парметра. В основу метода положен подход кластерного анализа - выявление характеризующих структур. Метод претендует на наиболее полный охват различительной информации, содержащейся в последовательности компонент вектора прохождения, как то:

- а) количество энерговыделения в каждом ряду;
- б) форма функции распределения энерговыделений по длине установки.

Ереванский физический институт

Ереван 1986

Preprint EOM-940(9I)-86

S.Kh. ARUTUNYAN, A.A. CHILINGARYAN, R.P. KOKOULIN

ON NONPARAMETRICAL METHODS OF SUPPRESSION OF
NUCLEAR CASCADE BACKGROUND IN MEASURING MUON SPECTRUM
WITH A PAIRMETER

A new nonparametrical method to identify events of nuclear muon interaction is suggested in the problem of measuring muon spectrum by means of a multi-row detecting device of a pair-meter type. The method is based on the cluster analysis approach - reveal of characterizing structures. The method pretends on the most entire scope of distinctive information available in a succession of passage vector components, such as: a) amount of energy release in each row; b) shape of energy release distribution function along the device's length.

Yerevan Physics Institute

Yerevan 1986

На сегодняшний день можно считать, по-видимому, доказанным, что наиболее подходящей установкой для исследования энергетического спектра мюонов космического излучения с энергией вплоть до 100 ТэВ является установка типа парметра - периодическая система детекторов ионизации, использующая энергетическую зависимость прямого рождения электрон-позитронных пар [1]. Единственная работающая установка этого типа была запущена в Японии, хотя разработке этой идеи были посвящены многочисленные теоретические работы [2] .

В связи с возможностью сооружения установки типа парметра в комплексе АНИ была проведена настоящая работа по исследованию возможности подавления фона от ядерных взаимодействий мюонов при измерении спектра мюонов с энергиями вплоть до 100 ТэВ.

Из-за значительных флуктуаций в продольном развитии каскады, вызванные ядерными взаимодействиями (так называемые "ядерные" каскады) низкоэнергичных мюонов, имитируют процессы множественных взаимодействий с малыми передачами мюонов гораздо

больших энергий. Предложенный в работе [3] алгоритм дискриминации, основанный на разнице в распределении расстояний между рядами, в которых были зарегистрированы самые большие толчки, позволил продолжить область энергий, эффективно исследуемых установкой параметр до 5 ТэВ.

Нами показана принципиальная возможность достижения более высокого уровня дискриминации с помощью методов, развитых в рамках подхода распознавания образов. Исследования по проверке предлагаемых методов проводились на материале имитационного моделирования. Программой моделирования отклика периодической системы детекторов ионизации на прохождение мюонов высокой энергии [3] были получены наборы, соответствующие разным типам установок (варьировалось число рядов, вещество поглотителя) и разным критериям отбора событий. Изменением значений параметров программы можно было "выключать" ядерные взаимодействия. Эти наборы - обучающие выборки - записывались в базы данных центра данных АНИ [4].

Рассмотрим выборочные матрицы корреляции случаев прохождений с "чисто" электромагнитным (рис.1) и "чисто" ядерным (рис.2) взаимодействиями. Ромбиками отмечены значения корреляции

$\rho \geq 0,4$, крестиками - $\rho < 0,4$. Для "ядерных" событий характерны большие корреляции между значениями откликов в соседних рядах. Эти каскады более энергичны и компактны. На рис.3 и 4 приведены примеры каскадов ядерных и электромагнитных. Ядерным каскадам, в отличие от электромагнитных, присущи определенные структура и форма, что является ценной различительной информацией. Наиболее полное использование этой информации являлось критерием в разработке и отборе методик иденти-

фикации типа прохождения. В данной работе предлагаются два метода кластерного анализа и метод КЭС (в метрике Евклида, что обусловлено топологией признакового пространства событий). Параметры методов имеют феноменологическую, опытную природу и зависят от значений критериев отбора и оценок энергии. Как показано в [3], простым и эффективным методом оценивания энергии в установках типа парметра является использование порядковых статистик. В широком диапазоне энергий (3-100 ТэВ), для установок с 20 - 30 рядами, логарифм 5 - 10-й порядковой статистики убывающего вариационного ряда энерговыделений в детекторах пропорционален логарифму энергии мюона, прошедшего через установку. Таким образом, для отбора событий с энергией выше определенного порога достаточно потребовать, например, чтобы значение 5-й порядковой статистики было больше заданной величины.

Кластерный анализ события начинается с процедуры формирования кластера. Кластером назовем последовательную группу значений энерговыделений в рядах установки, содержащую 5 значений (наибольшее обязательно среди них) больше порогового. После этапа формирования вырабатывается критерий кластеризации.

Первый из предложенных нами кластерных характеристик - длина кластера - определяется разностью крайних номеров детекторов, попавших в кластерную группу. Если длина кластера больше 8, это прохождение классифицируется как электромагнитное. В противном случае событие идентифицируется как ядерное и обрабатывается как фоновое. На рис. 5 и 6 представлены гистограммы величины $\ln [E_\mu / n_5]$, где E_μ - энергия мюона, n_5 - значение 5-й порядковой статис-

тики вариационного ряда энерговыделений до (рис.5) и после (рис.6) классификации выборки, составленной из смеси ядерных и электромагнитных событий. Из них 50 - ядерных и 40 - электромагнитных в соответствии с реальной фоновой ситуацией. Видно, что алгоритм эффективно выделяет левую - "ядерную" часть распределения.

I. Метод характерных расстояний

Рассмотрим совокупность $\{a_i\}$, $i = 1, n$ откликов на прохождение мюона через n - рядную установку. На этой совокупности решим следующую задачу. Определим критерий идентификации вида прохождения, соответствующий наиболее полному использованию различительной информации ядерного и электромагнитного (ЭМ) прохождений. Проведем процедуру формирования кластера на совокупности $\{a_i\}$ и сделаем следующий отбор, продиктованный условиями физической задачи

$$a'_i = \begin{cases} 1, & \text{если } a_i < 100 \\ a_i, & \text{если } a_i \geq 100 \end{cases} \quad (100 - E_\mu \sim 5 \text{ ТэВ})$$

Совокупность $\{a'_i\}$ порождает последовательность номеров $\{j_i\}$, для которых $a_{j_i} \geq 100$. Пусть число членов последовательности равно n^* .

Применим преобразование инверсии:

$$a_i^* = \frac{A}{a_i}, \quad \text{где } A^{-1} = n^{-1} \cdot (n - n^* + \sum_{i=j_1}^{n^*} \frac{1}{a_i}).$$

Преобразование инверсии, унифицируя ЭМ и ядерные прохождения,

дает возможность более строго и корректно использовать формализм математической статистики в задаче идентификации по форме каскада. По физическим условиям задачи прохождение характеризуется пятью значениями $\alpha_i \geq 100$. Рассмотрим вектор $\vec{\alpha}^* = (\alpha_j^*)$. Последовательность пяти номеров $\{j\}$ повторяет часть последовательности $\{J\}$, начиная с некоторого номера, так, чтобы выполнялись следующие условия:

- а) максимальное α_{\max}^* из совокупности $\{\alpha_i^*\}$ является компонентой вектора $\vec{\alpha}^*$;
- б) при условии а) разность $(j_5 - j_1)$ имеет минимальное значение.

Итак, совокупность пятимерных векторов $\vec{\alpha}$ сосредоточена в n -мерной сфере единичного радиуса. Для компонент векторов, соответствующих ядерному прохождению, свойствен:

- а) порядок следования с некоторым монотонным убыванием значений α_i^* ;
- б) сравнительно малые значения разницы $j_5 - j_1$:

$$5 \leq j_5 - j_1 \leq 8.$$

Таким образом, векторы $\vec{\alpha}$ ядерного прохождения "нанизаны" на координатные оси. Последовательность компонент вектора ЭМ прохождения:

- а) лишена какой-либо систематики (монотонности);
- б) разбросана по всей координатной сетке

$$j_5 - j_1 > 8.$$

Введем весовой вектор $\vec{W} = (W_i)$, $i = 1, 5$, $\sum_{i=1}^5 W_i = 1$, который имитирует обратную форму каскада от ядерного прохождения.

Тогда

$$\vec{W} \cdot \vec{\alpha}_{яг}^* \sim 0 \text{ (так как } \cos \alpha \sim 0 \text{),}$$

$$\vec{W} \cdot \vec{\alpha}_{эм}^* \neq 0 \text{ (так как } \cos \alpha \neq 0 \text{).}$$

Пусть P_K - параметр кластеризации и

$$P_K = \sum_{i=j_1}^{j_5} \alpha_i^*$$

Тогда для $P'_K = \sum_{m=1,5}^{j_5} \alpha_i^* \cdot W_m$ получается:

$$P'_{Kяг} \ll P_{Kяг},$$

$$P'_{Kэм} \ll P_{Kэм},$$

т.е. достигается разделимость значения P_K от использования информации о форме каскада. Отметим, что $P_{Kяг} \sim P_{Kэм}$.

Дальнейшее улучшение разделимости значений P_K связано с использованием информации о степени разброса компонент $\vec{\alpha}^*$ по координатной сетке, т.е. с использованием значения разности $j_5 - j_1$. С этой целью введем новый фактор ℓ_m

$$P''_K = \sum_{m=1,5}^{j_5} \alpha_i^* \cdot W_m \cdot \ell_m,$$

где $\ell_m = \begin{cases} (j_m - j_{m-1})^2, & m \neq 1 \\ 1, & m = 1 \end{cases}$. Так как для ядерных прохождений $\ell_m \sim 1$, то

$$P''_{Kяг} \sim P'_{Kяг}.$$

Для ЭМ прохождений $\ell_m \gg 1$, поэтому

$$P''_{Kэм} \gg P'_{Kэм}.$$

На рис.7 представлено распределение величины P''_K для ис-

следуемой установки при энергиях мюонов $E_{\mu} \sim 5$ ТэВ. Ядерным событиям (они представлены крестиками) соответствуют значения $P_k'' < 0,05$. При таком пороге выделения ядерного фона число ошибочно классифицированных событий равно II. При увеличении значения порога улучшается эффективность действия классификатора, но резко увеличивается объем необходимой статистики, что связано с нежелательно большими затратами "машинного времени". 4-я строка таблицы представляет действие классификатора K - ближайших соседей (КБС) (метрика Евклида) на преобразованную выборку. Суть этого преобразования заключается в усилении структурной разницы отклика установки на прохождение с ядерным или электромагнитным взаимодействием: 30-мерный "вектор-отклик" сжимается в 2-мерный. Преобразованный "вектор" имеет сильно неравные компоненты, компоненты электромагнитного "вектора" примерно одинаковые. На двумерной плоскости (x, y) ядерные события оказываются сосредоточенными вдоль осей координат, электромагнитные события - вдоль прямой $y = x$. Такое разделение приводит к небольшим ошибкам КБС классификации: 4 - среди ядерных (потери в эффективности) и 10 - среди электромагнитных (потери в статистике).

Результаты

В таблице представлены численные результаты некоторых характеристик используемых методов до (I-я строка) и после классификации: кластерным методом длин (2-я строка) кластерным методом характерных расстояний (3-я строка) и методом КБС (4-я строка).

Таблица

	Число ошибок классифик.	\bar{y}	$\sigma_{\bar{y}}$	η	$S\Omega_{ef}$
До классификации		-1,8	0,37	0,054	
Метод длины кластера	7 яд.+ 2 ЭМ	-1,3	0,2	0,6	0,12
Метод характерного расстояния	5 яд.+ 6 ЭМ	-1,25	0,2	0,71	0,099
Метод КБС	4 яд.+ 10 ЭМ	-1,31	0,22	0,5	0,086

Здесь \bar{y} - выборочное среднее значение величины $y = \ln(E_{\mu}/n_5)$;
 $\sigma_{\bar{y}}$ - выборочная дисперсия величины y ;
 η - нормированная эффективность классификатора

$$\eta = \frac{N_{\alpha}}{N_0} \exp(\gamma(\bar{y}_{\alpha} - \bar{y}_0)),$$

где N_0 - истинное число событий типа " α ", N_{α} - число отображенных данным методом событий типа " α ". Оптимальный критерий классификации должен максимизировать величину η . Из таблицы видно, что наибольшее значение η достигается методом характерного расстояния .

$S\Omega_{ef}$ - "эффективная" светосила установки:

$$S\Omega_{ef} = S\Omega_0 \cdot \exp(-\gamma^2 \delta^2 / 2),$$

где $\delta = 1/\sqrt{\alpha T}$ - дисперсия логарифма оценки энергии [5,6],
 $\alpha = \frac{1}{137}$, T - толщина установки в каскадных единицах.

Для рассматриваемой установки (30 рядов железа, $T = 300$ р.е.)

$\delta = 0,68$, и для идеальной установки получается:

$$S\Omega_0 = 0,19.$$

Таким образом, предложенные методы выделения ядерного фона обеспечивают эффективную светосилу, составляющую от 45 (метод КБС) до 60% (метод длина кластера) эффективной светосилы идеального случая. Значение величины \bar{C} для выборки электромагнитных событий равно 1,2.

Сравнение значений величины \bar{C} из таблицы до и после классификации показывает возможность почти полной дискриминации ядерного фона, а затем и продвижения в область больших исследуемых энергий.

В заключение авторы выражают благодарность Ерлыкину А.Д. за полезное обсуждение и ценные замечания.

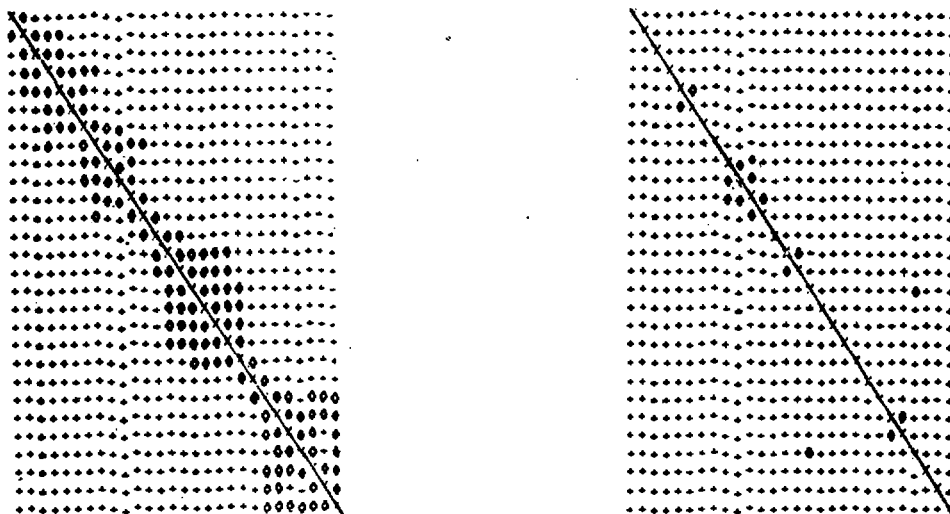


Рис. 1,2 Корреляционные матрицы, соответствующие ядерным (рис.1) и электромагнитным (рис.2) взаимодействиям:

◇ - корреляции, превышающие 0,4

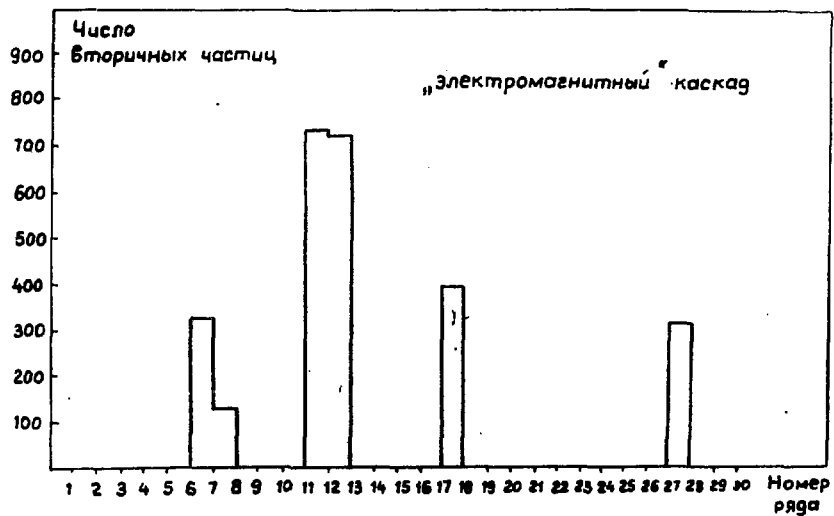
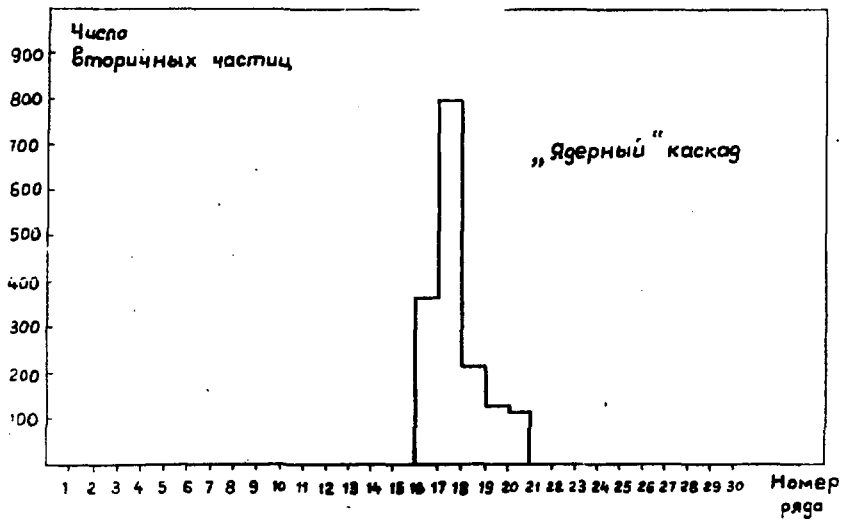


Рис.3,4 Распределения числа вторичных частиц, родившихся в результате ядерного (рис.3) и электромагнитного (рис.4) взаимодействий мюона, прошедшего через 30 рядов установки

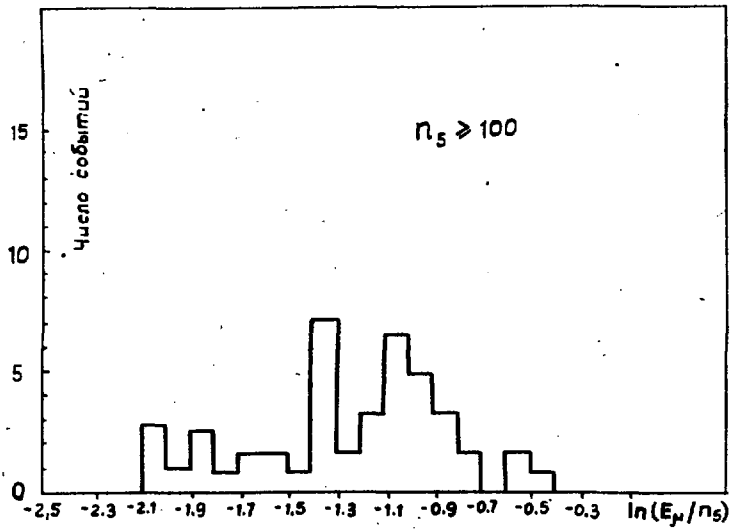
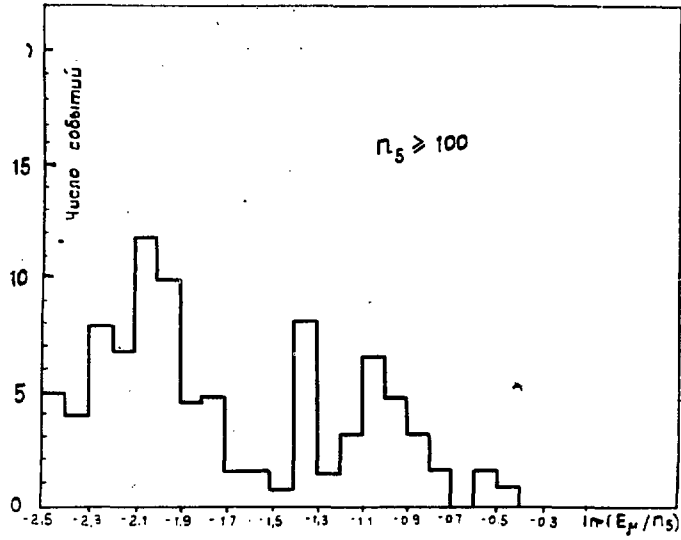


Рис.5,6 Гистограммы величины $\ln[E_\mu/n_5]$ до (рис.5) и после (рис.6) классификации

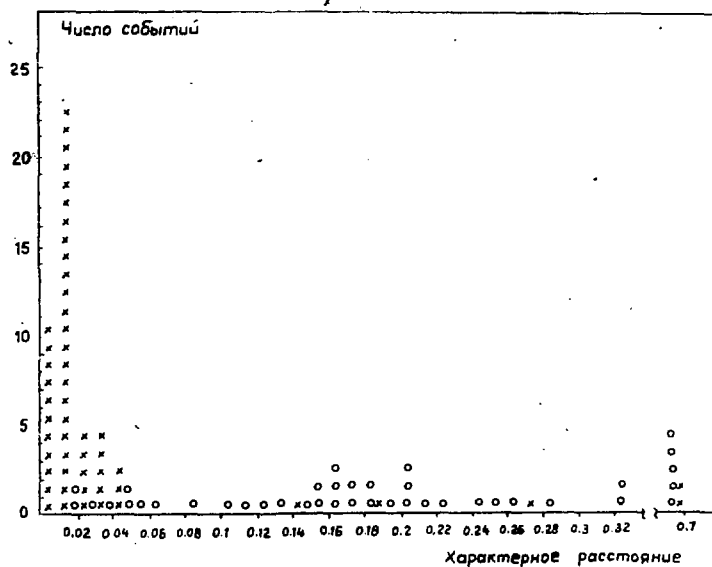


Рис.7 Гистограмма величины R''_k - характерного расстояния:

× - ядерные события;

○ - электромагнитные события

СПИСОК ЛИТЕРАТУРЫ

1. Алексеев И.С., Зацепин Г.Т. Тр. Моск. Межд. конф. по космическим лучам, 1960, т.1, с.324.
2. Mitsui K., Okada A., Kitamura T. et al. A calorimeter to study cosmic ray muon interaction, ICR-Report-79-79-3, Tokyo Univ., 1979.
3. Kokoulin R.P., Monte-Carlo calculation for Pair-Meter, ICR-Report-102-82-5, Tokyo Univ., 1982.
4. Akopov N.Z., Harutjunyan Sv. Kh., Chilingaryan A.A. The design principle and structure of "ANI" data centre, EPHI- 819(46)-85, Yerevan, 1985.
5. Kokoulin R.P., Petrukhin A.A. Comparative analysis of possibilities of various techniques muon spectrum measurements, 18th ICRC, Bangalore, 7,17-20, 1983.
6. Kokoulin R.P. Possibilities of multi-TeV cosmic ray muon spectrum measurements by using large scale Pair-Meter detectore, ICR-Report-103-82-6, Tokyo Univ. 1982.

Рукопись поступила 8 сентября 1986 г.

С.Х.АРУТЮНЯН, Р.П.КОКОУЛИН, А.А.ЧИЛИНГАРЯН

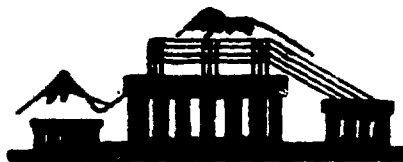
**О НЕПАРАМЕТРИЧЕСКИХ МЕТОДАХ ПОДАВЛЕНИЯ ФОНА ЯДЕРНЫХ
КАСКАДОВ ПРИ ИЗМЕРЕНИИ СПЕКТРА МЮОНОВ С ПОМОЩЬЮ УСТАНОВКИ
ТИПА ПАРМЕТРА**

**Редактор Л.П.Мукаян
Технический редактор А.С.Абрамян**

**Подписано в печать 19/ХП-86г. ВФ-07493 Формат 60x84/16
Офсетная печать. Уч.изд.л. 1,0 Тираж 299 экз. Ц. 15 к.
Зак.тип. 659 Индекс 3624**

**Отпечатано в Ереванском физическом институте
Ереван 36, Маркаряна 2**

индекс 3624



ЕРЕВАНСКИЙ ФИЗИЧЕСКИЙ ИНСТИТУТ