

О массовом составе при ультравысоких энергиях по открытым данным ОЖЕ коллаборации (статистический подход с использованием функции копула)

Author: Dr. КИРИЛЛОВ, Александр¹ НИИЯФ МГУ Москва 119991 Ленинские горы 1
Co-Author: Dr. КИРИЛЛОВ, Илья² Концерн "Энергоатом" Москва 109507, Ферганская 2
Corresponding Author: krl@dec1.sinp.msu.ru

Аннотация

Для исследования динамики изменения массового состава космических лучей самых высоких энергий используются данные ($3 \cdot 10^{18}$ -- $5 \cdot 10^{19}$) эВ, предоставленные Оже-коллаборацией для всеобщего использования, по состоянию на 01.01.2014 (1% всех имеющихся у коллаборации данных). Исследование статистическими методами с привлечением аппарата описания внутренних связей многомерных распределений (функции копула) показывает утяжеление массового состава за счёт обогащения более тяжёлой компонентой группы средних ядер. Это утяжеление имеет сложный характер и происходит постепенно с ростом энергии. Резких, существенных изменений массового состава данные ($3 \cdot 10^{18}$ -- $3 \cdot 10^{19}$) эВ не показывают.

1. Введение

Исследование массового состава излучения ультра высоких энергий ($>1 \text{EэВ} = 1 \cdot 10^{18} \text{эВ}$) – трудная задача, т.к. каскадные флуктуации не позволяют построить модель ШАЛ, которая может указывать на значение массы (или заряда) первичной частицы индивидуального ливня на базе данных конкретной экспериментальной установки достаточно определённо. Поэтому естественно использовать статистический подход, который параметрами функций распределения может восполнить недостающие описательные возможности эксперимента. Наиболее статистически богатой является информация от «поверхностных» детекторов (Surface Detectors) коллаборации Оже [1]. Россия не участвует в этой коллаборации, но 1% всех имеющихся у коллаборации данных предоставляется для всеобщего использования. Ковёр ШАЛ покрыт этими детекторами триангулярно с шагом 1,6 км. Ливень регистрируется путём записи показаний детекторов. Эти записи, дополненные указанием координат детекторов, координат оси ливня, энергии ливня, времени и направления его прихода определяют событие. Всего в интервале энергий (1.8 – 50.0) EэВ на 01.01.2014 дано для всеобщего использования 2417 событий.

По причине актуальности и противоречивости имеющихся данных и их интерпретации в добавление к известному [2,3] подходу оценки массового состава, основанному на значении определяющего параметра X_{\max} - глубины максимума ливня, в последнее время рассмотрен ряд физических задач и предложены методы основанные, в частности, на зависимости от массового состава регистрируемой мюонной компоненты [4,5], пространственного распределения [6], временного развития ливня [7,8]. Результаты по исследованию временных параметров в настоящей работе не используются.

В настоящей работе рассматриваются 2 определяющих переменных параметра: геометрическая (r_i) – расстояние i -того детектора от оси ливня (метры) и энергетическая (e_i) – энергия зарегистрированная детектором (VEM) $i = 1 \div n$. На этих переменных с учётом вышеприведённых результатов [2 – 6] строятся и анализируются пары функций, идентифицирующих данное событие. Естественно в первую очередь рассмотреть средние: каждое событие идентифицировать значением среднего расстояния откликнувшихся детекторов от оси и значением средней энергии зафиксированной этими детекторами. При этом очевидно, что при фиксированной энергии рост $\langle r_i \rangle$ указывает на утяжеление массового состава. Также очевидно, что при постоянном массовом составе рост $\langle r_i \rangle$ указывает на рост энергии усредняемых событий. Перераспределение значений

энергетических показаний детекторов, вызванное изменением массового состава, приводит к изменению формы функции энергетического распределения и отражается на значении средних.

В работе на последовательности энергетических интервалов с ростом энергии исследуется эволюция характеристик 2-мерных функций распределения, описывающих совокупность событий в каждом интервале. Рассматривается и эволюция маргиналов (частных распределений) и эволюция их функциональных связей. Последнее использует известные меры функции копулы, которая описывает внутренние функциональные связи многомерных распределений. Никаких предположений о виде непрерывных функций распределения (часто неявных, например, нормальность какого-либо распределения или независимость частных распределений) не делается; предполагается только неподвижность 1% выборки, предоставленной для всеобщего использования.

2. Последовательность энергетических интервалов

Область «лодыжки» ((1,8-- 50) ЕэВ, $E_{\text{анкл}}=4$ (ЕэВ) [9]) разделена на 12 энергетических интервалов так, чтобы при примерном равенстве суммарного количества сработавших детекторов (1071 – 1076) каждый интервал содержал целое число событий (ливней), что определяет их нижние границы: (1.8, 1.915, 2.013, 2.157, 2.322, 2.527, 2.807, 3.154, 3.609, 4.292, 5.729, 10.0) ЕэВ (см. стрелки на рис.1). Результаты статистической обработки интервала ассоциируются с его нижней границей. Из-за роста длины интервалов, убывания количества событий и соответствующего роста статистических погрешностей такая «ассоциация» вносит неточности, поэтому рассмотрение самого энергичного интервала ((1- 50) ЕэВ содержащего всего 106 самых энергичных событий) дополняется неформальным рассмотрением в Дискуссии.

Естественно с ростом энергии из-за крутого энергетического спектра $\sim E^{-3}$ увеличивается длина этих интервалов, а количество зарегистрированных событий в интервале убывает (более энергичные ливни регистрируются большим количеством детекторов). Эволюцию этих характеристик можно видеть из рис. 1.

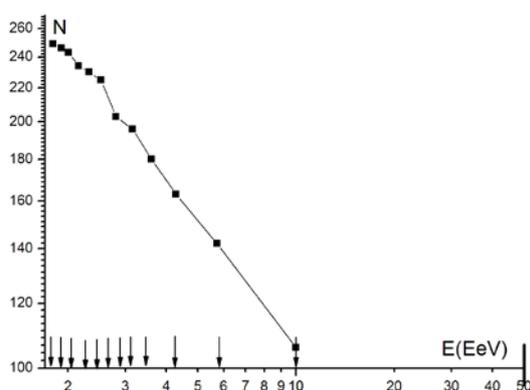


Рис. 1. Энергетические интервалы (нижние границы даны стрелками) и количество событий в каждом интервале.

Рис.1, показывает: степенное убывание количество событий (прямая в двойных логарифмических координатах); излом при $E=(2.5 - 2.8)$ ЕэВ; рост наклона после излома.

Количество событий в энергетическом интервале обратно «пропорционально» среднему количеству детекторов регистрирующих событие из этого интервала и может

рассматриваться как мера ширины ливня, которая в свою очередь может рассматриваться как мера «тяжёлости» массового состава. Степенной характер убывания определяется степенным спектром первичного потока, наклон – естественным уширением ливней при возрастании первичной энергии.

Наблюдаемый на рис.1 излом – излом спектра и массового состава в области «лодыжки». Изменение наклона – утяжеление массового состава. Это утяжеление может иметь различные причины: дефицит лёгкой компоненты или обогащение тяжёлой компоненты или утяжеление средней компоненты или различные комбинации этих основных причин. При данной статистике эта характеристика идентифицировать причину более точно не может.

Рассмотрение других характеристик разбиения, отнесённых к каждому энергетическому интервалу (количество детекторов, зарегистрировавших ливень; расстояние от оси ливня; зарегистрированная энергия) не противоречит изложенному, а подтверждает его; но существенно не дополняет.

3. Многомерные распределения

а) эволюция маргиналов

Как сказано выше, для каждого ливня в каждом энергетическом интервале рассчитывались пары идентифицирующих функций и исследовалась эволюция параметров этих 2-мерных функций с ростом энергии интервалов.

На рис. 2 представлены средние двумерные функции (см. 1) в разделе Дискуссия) для 3 интервалов: (3.16 – 3.64; 4.92 – 7.29; 10.0 – 50.0) ЕэВ.

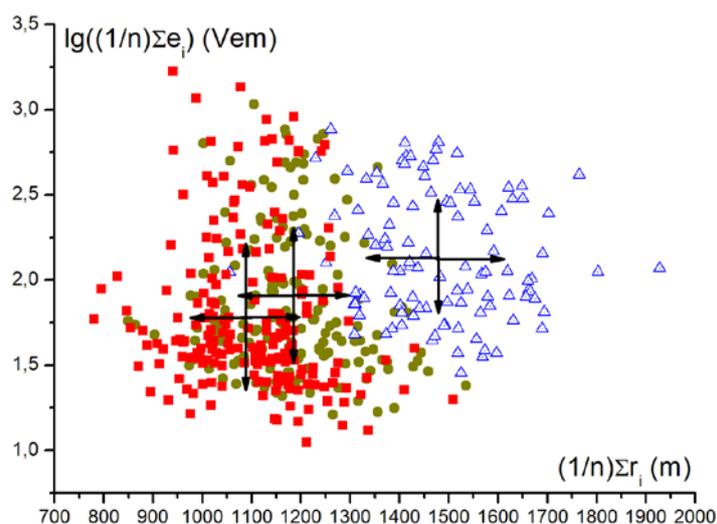


Рис. 2. Идентифицирующие функции средних значений расстояний - $\langle r_i \rangle$ и логарифмов энергий - $\langle \lg e_i \rangle$ сработавших (n) детекторов. Квадрат; круг; треугольник - соответственно для интервалов (3.16 – 3.64; 4.92 – 7.29; 10.0 – 50.0) ЕэВ. Кресты – центры тяжести распределений с указанием стандартных отклонений по осям.

Из рис. 2 видно, что с ростом энергии (квадрат; круг; треугольник) распределения сдвигаются вправо вверх и становятся более компактными. Это естественная картина, обусловленная ростом первичной энергии. Рост среднего расстояния от оси обусловлен добавлением периферических детекторов. Рост средней энергии не противоречит утяжелению спектра; величина и скорость этого роста на имеющейся статистике не показывают значительных качественных скачков. Величина и скорость изменения верхних и нижних границ идентифицирующих функций не показывают значительных качественных изменений массового состава.

Можно также заметить, что в наименее энергичном интервале (квадраты) поле точек показывает два размытых кластера с координатами центров (1020,1.6) и (1150,1.5).

(Можно грубо интерпретировать, что левый кластер соответствует лёгкой компоненте, а правый – тяжёлой.) С ростом энергии (круги, треугольники) эти кластеры продолжают размываться и в наиболее энергичном интервале начинает выделяться средняя часть вблизи центра тяжести.

Более определённые результаты может показать исследование маргиналов, приведённых на рис. 3.

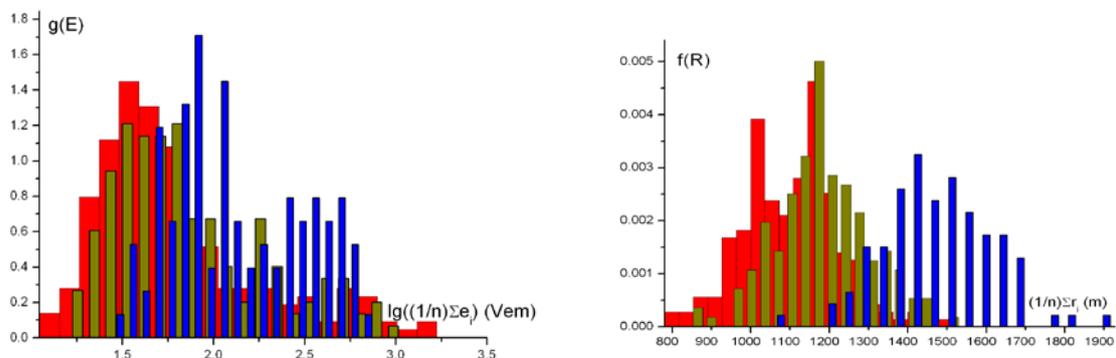


Рис. 3. Плотности функций распределения средних расстояний (слева) и логарифмов средних значений энергий (справа) сработавших детекторов для каждого зарегистрированного события. (широкие столбцы; более узкие; самые узкие – соответственно для интервалов (3.16 – 3.64; 4.92 – 7.29; 10.0 – 50.0) ЕэВ).

Здесь из левой части рисунка ясно виден отмеченный выше сдвиг средних значений этих распределений вместе с уменьшением их области определения. Изменение формы этих распределений сводится к тому, что в наиболее энергичной области распределение становится более плоским (приближается к равномерному), т.е. при наличии каскадов с малыми значениями $\sum r_i$ (лёгкая компонента), так и с большими (тяжёлая компонента) увеличивается доля событий соответствующих средней компоненте. Такое явление (увеличение доли средней компоненты) должно сопровождаться увеличением количества событий с энергией в области средних значений. Именно это и видно из правой части. Изменение формы энергетического распределения сводится к его уширению в области среднего. Заметим также, что для самого энергичного интервала отсутствие ливней с очень высокими показателями средней энергии нельзя рассматривать как отсутствие лёгкой компоненты, так как с ростом энергии ливень (даже фотонный) регистрируется большим количеством детекторов и явно выраженный энергетический пик (характерный для лёгкой компоненты) сглаживается усреднением. Также усреднением при имеющейся статистике может быть объяснено отсутствие ливней с очень низкими показателями средней энергии, характерными для тяжёлой компоненты.

Отмеченную выше размывку кластеров с ростом энергии интервала можно ясно видеть в эволюции формы радиального распределения. Два выраженных пика в наименее энергичном интервале сливаются, а затем размываются вблизи средних значений самого энергичного интервала.

Итак, анализ эволюции маргиналов (уплощение радиального и уширение энергетического распределений) указывает, что утяжеление массового состава обусловлено ростом доли средней компоненты.

б) внутренние связи частных распределений

Многомерное распределение определяется своими маргиналами и внутренними связями между частными распределениями, определяемые $C(u, v)$ - копула функцией, где u, v – обратные функции маргиналов [10]. {*}{Из-за ограниченности объёма сообщения сноски с определением копула функции, теорема Скляра, построение эмпирической и частотной копул опущены. Это можно найти например в Википедии.} Простейшая

(линейная) связь (соответствие большим значениям одной переменной большим значениям другой, и аналогично с малыми значениями) описывается γ - общеизвестным коэффициентом корреляции Пирсона, который в настоящей задаче мал и непоказателен, что можно видеть из рис. 2, как основное линейное направление поля точек, определяемое двумерным распределением. Вместе с тем трудно предполагать независимость геометрической и энергетической переменных, поэтому необходимо рассмотреть нелинейные связи, определяемые копулой и выражаемые её мерами.

Напомним построение и интерпретацию наиболее употребительных мер – Кендала, ρ – Спирмана, γ – Гини. Пусть (X_1, Y_1) и (X_2, Y_2) независимые и одинаково распределённые пары векторов определяется как вероятность соответствия минус вероятность несоответствия: $\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$.

Значения мер τ , ρ , γ выражаются через копулу:

$$\tau = 4 \iint_{I^2} C(u, v) dC(u, v) - 1 = 4E(C(u, v)) - 1,$$

$$\rho = 12 \iint_{I^2} uv dC(u, v) - 3 = 12 \iint_{I^2} (C(u, v) - uv) dudv,$$

$$\gamma = 4 \left(\int_0^1 C(u, 1-u) du - \int_0^1 (u - C(u, u)) du \right).$$

Эти выражения показывают варианты интерпретации мер. Интерпретация τ также одержит-

ся в его определении. Т.к. $C(u, v) = uv$ - копула независимых распределений, ρ может интерпретироваться как разность объёмов рассматриваемой и независимой копул. γ интерпретируется как разность площадей сечений по побочной диагонали области определения копулы ($I^2 = \{0,1\} \times \{0,1\}$) и разности сечений по главной диагонали максимальной и рассматриваемой копул.

Практически N значений многомерной идентифицирующей функции определяют эмпирическую $C_N(i/N, j/N)$ и частотные $c_N(i/N, j/N)$ копулы – аналоги эмпирической функции распределения и гистограммы в одномерном случае.

$$\tau = \frac{2N}{N-1} \sum_{i=2}^N \sum_{j=2}^N \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} (c(\frac{i}{N}, \frac{j}{N})c(\frac{p}{N}, \frac{q}{N}) - c(\frac{i}{N}, \frac{q}{N})c(\frac{p}{N}, \frac{j}{N}))$$

$$\rho = \frac{12}{N^2-1} \sum_{i=1}^N \sum_{j=1}^N (C(\frac{i}{N}, \frac{j}{N}) - \frac{i}{N} \frac{j}{N})$$

$$\gamma = \frac{2N}{[N^2/2]} \left(\sum_{i=1}^{N-1} C(\frac{i}{N}, 1 - \frac{i}{N}) - \sum_{i=1}^N (\frac{i}{N} - C(\frac{i}{N}, \frac{i}{N})) \right).$$

Анализ многомерных распределений (в частности графических образов) труднее, чем в одномерном случае, но выполняется в основном также: сравнением с известными базовыми случаями. В общем случае используются: минимальная копула – прямая треугольная пирамида с основанием $u + v \geq 1$ на квадрате $\{0,1\} \times \{0,1\}$ с вершиной в $(1,1,1)$; независимая копула – четверть седла ($uv \leq 1$); максимальная копула – прямая четырёхугольная пирамида на квадрате $\{0,1\} \times \{0,1\}$ с вершиной в $(1,1,1)$. Все меры минимальной, независимой и максимальной копул равны соответственно -1, 0, +1.

Интересна область наиболее высоких энергий.

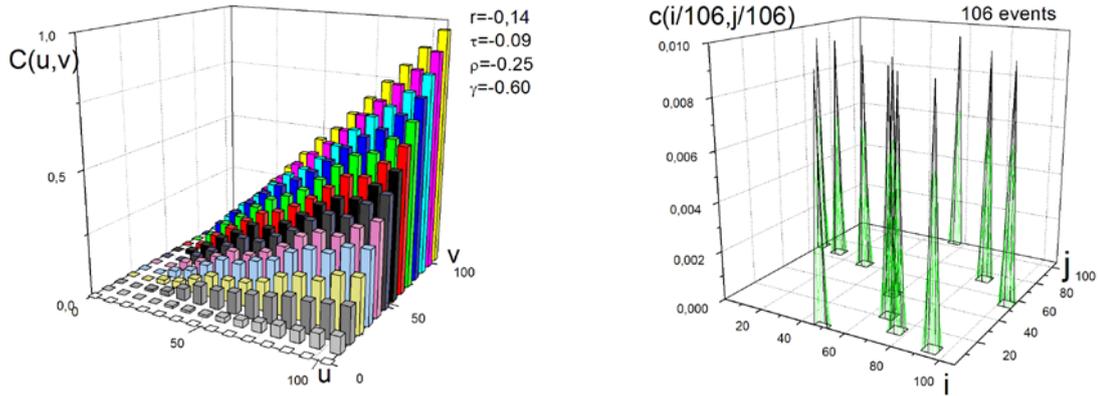


Рис. 4. Эмпирическая (слева) и частотная (справа) копулы функции распределения $\langle r_i \rangle$ - средних значений расстояний сработавших детекторов и $\lg \langle e_i \rangle$ - логарифмов средних зарегистрированных ими энергий от ливней в интервале (10.0 – 50.0) ЕэВ.

Из рисунка видно, что рассматриваемая копула несколько ($\rho = -0.25$) ниже четверти седла (независимой копулы), причём в области побочной диагонали ($u + v = 1$) наблюдается некоторое возвышение особенно заметное у концов диагонали. Количественно это подтверждается приведёнными значениями мер. Частотная копула явно показывает обеднение на главной диагонали и обогащение на побочной, что подтверждается значением ($|\gamma| > 0.5$) меры Гини.

Вспоминая грубую интерпретационную модель: лёгкая компонента – малые r (соответственно малые u и i) и большие e (соответственно большие u и i), тяжёлая – всё наоборот, можно понимать возвышение у концов побочной диагонали как присутствие лёгкой и тяжёлой компонент, а возвышение в её центральной части – доля средней компоненты, что, как главное, и отмечается наибольшим значением из рассматриваемых мер – мерой Гини γ . Это рельефно показывается сливающимися столбцами частотной копулы и практически нулевой мерой Кендала τ .

Рассмотрение эволюции внутренних связей не показывает резких существенных изменений в области $5 \cdot 10^{18}$ -- $5 \cdot 10^{19}$ (ЭВ). Всё это подтверждает полученное выше и из анализа эволюции маргиналов указание на утяжеление массового состава в основном за счёт увеличения доли средней компоненты.

4. Дискуссия

Наряду с изложенным случаем средних значений расстояний и логарифмов средних зафиксированных энергий рассматривались также и другие варианты пар идентифицирующих функций:

$$\begin{aligned}
 1) \quad f_x &= \frac{1}{n} \sum_{i=1}^n r_i & f_y &= \lg\left(\frac{1}{n} \sum_{i=1}^n e_i\right) & 6) \quad f_x &= r_{\max} - r_{\min} & f_y &= E_{r_{\min}} - E_{r_{\max}} \\
 2) \quad f_x &= \frac{1}{n} \sum_{i=1}^n r_i & f_y &= \lg\left(\frac{1}{n} \sum_{i=1}^n e_i r_i\right) & 7) \quad f_x &= r_{\max} - r_{\min} & f_y &= E^0 / \sum_{i=1}^n e_i \\
 3) \quad f_x &= \frac{1}{n} \sum_{i=1}^n r_i & f_y &= \lg\left(\frac{1}{n} \sum_{i=1}^n \frac{e_i}{r_i}\right) & 8) \quad f_x &= \frac{r_{\max} - r_{\min}}{\sum_{i=1}^n r_i} & f_y &= \frac{E_{r_{\min}} - E_{r_{\max}}}{\sum_{i=1}^n e_i} \\
 4) \quad f_x &= \sum_{i=1}^n r_i e_i / E^0 & f_y &= \lg \sum_{i=1}^n \frac{e_i}{r_i} & 9) \quad f_2 &= \frac{r_{cw} - r_{\min}}{r_{\max} - r_{\min}} & f_y &= \frac{E_{r_{\min}} - E_{r_{\max}}}{\sum_{i=1}^n e_i}, \\
 5) \quad f_x &= \frac{1}{n} \sum_{i=1}^n r_i & f_y &= \frac{1}{n} \sum_{i=1}^n e_i
 \end{aligned}$$

где: r_i – расстояние i -того детектора от оси ливня (метры), e_i – энергия зарегистрированная детектором (VEM), n – количество детекторов, зарегистрировавших событие, $i = 1 \div n$, r_{\min} , (r_{\max}) – расстояние минимально (и соответственно максимально) удалённого детектора от оси ливня, $r_{\text{св}}$ – среднее расстояние от оси ливня детекторов, зарегистрировавших данный ливень, $E_{\text{T min}}$, ($E_{\text{T max}}$) – энергия, показанная соответственно r_{\min} , (r_{\max}) детекторами, E^0 – энергия первичной частицы ливня.

Это рассмотрение подтвердило изложенное выше, более рельефно показало отдельные моменты, но авторам не удалось аргументировано выявить дополнительные существенные закономерности.

Во Введении говорилось о необходимости дополнить статистическое рассмотрение самого энергичного интервала 10 – 50 (ЕэВ) неформальным рассмотрением, поскольку рис. 1 показывает большой рост энергетического интервала, что может указывать на укрупнение спектра и на существенное облегчение массового состава. (Напомним, что величина энергетического интервала определяется суммарным количеством $1(070)$ сработавших детекторов.) По последним данным коллаборации Оже [10] укрупнение спектра начинается при 25 (ЕэВ) ($\lg E_{\text{break}}(\text{эВ}) = 19.41 \pm 0.2$ [9]), что не исключает также и второй возможности. Непосредственное рассмотрение показывает, что при $E < 20$ (ЕэВ) характеристики 78 ливней не противоречат вышерассмотренной эволюционной картине. При $E > 20$ (ЕэВ) из оставшихся 28 ливней (22 в промежутке (20 -30) ЕэВ и 6 при (30 – 50) ЕэВ) почти четверть можно предполагать тяжёлыми, что не позволяет ни предполагать, ни отвергать возможность облегчения массового состава при этих энергиях.

Авторы надеются, что не только расчёты, основанные на глубине максимума каскада, но увеличение статистики и использование (или разработка новых) более адекватных данной физической проблеме идентифицирующих функций и мер копулы позволят получить более точные результаты.

5. Заключение

Скромная статистика (1% всех имеющихся у коллаборации ОЖЕ данных, 106 событий в самом высокоэнергичном интервале $(1 - 5) \cdot 10^{19}$ ЭВ) заставляет ограничиться только качественным уровнем исследования.

Резких, существенных изменений массового состава имеющиеся данные (с учётом отмеченного выше о недостаточности доступной статистики при $E^0 > 2 \cdot 10^{19}$ (эВ)) в интервале $5 \cdot 10^{18}$ -- $5 \cdot 10^{19}$ (эВ) не показывают, но указывают на утяжеление массового состава. Это утяжеление развивается постепенно с ростом энергии, имеет сложный характер, и происходит в основном за счёт обогащения более тяжёлой компонентой группы средних ядер.

Литература

- [1] Pierre Auger Observatory Public Event Explorer <http://auger.colostate.edu/ED/>
- [2] Interpretation of the Depths of Maximum of Extensive Air Showers Measured by the Pierre Auger Observatory, JCAP 02 (2013) 026; arXiv:1301.6637
- [3] J. Matthews, A Heitler model of extensive air showers, Astropart. Phys. 22 (2005) 387
- [4] Precise determination of muon and electromagnetic shower contents from shower universality property. A. Yuskov, M. Ambrosio, et all, 32ND ICRC, BEIJING 2011 v2, 0687 p 75-78.
- [5] Applications of S_{μ}/S_{em} showers universality for mass composition and hadronic interactions studies. D.D.Urso, M. Ambrosio, et. all, 32ND ICRC, BEIJING 2011 v2, 0694 p 79-82.
- [6] Estimating primary mass composition of cosmic rays using geomagnetic spectroscopy. Pres.by J-N. CAPDEVIELLE. 33rd ICRC, RIO 2013 0770
- [7] Observations of the longitudinal development of extensive air showers with the surface detectors of the Pierre Auger Observatory. Pres. by GARCIA-GAMEZ. 33rd ICRC, RIO 2013 0694

- [8] Method for the primary mass composition study of ultra-high-energy cosmic rays with the Telescope Array surface detector. Pres. by S. Troitsky. 33rd ICRC, RIO 2013 0536
- [9] Measurement of the cosmic ray energy spectrum using hybrid events of the Pierre Auger Observatory, Eur. Phys. J. Plus 127 (2012) 87; arXiv:1208.6574
- [10] R.B. Nelsen. An Introduction to Copulas. Second Edition. Springer. 2006, 270p.