



Применение методов машинного обучения в Baikal-GVD: подавление фонового шума и выделение нейтринных событий

Мацейко Альберт^{1,2}, 5 курс МФТИ
Под руководством Харука И.В.

¹Московский физико-технический институт
(национальный исследовательский университет)

²Институт ядерных исследований РАН

Москва 2023

Исследование выполнено за счет гранта
Российского научного фонда 22-22-20063.

План доклада

1. Введение

- а) эксперимент Baikal-GVD
- б) μ и ν индуцированные события

2. Нейронные сети для:

- а) подавления шумовых срабатываний детекторов
- б) разделения нейтринных и мюонных событий

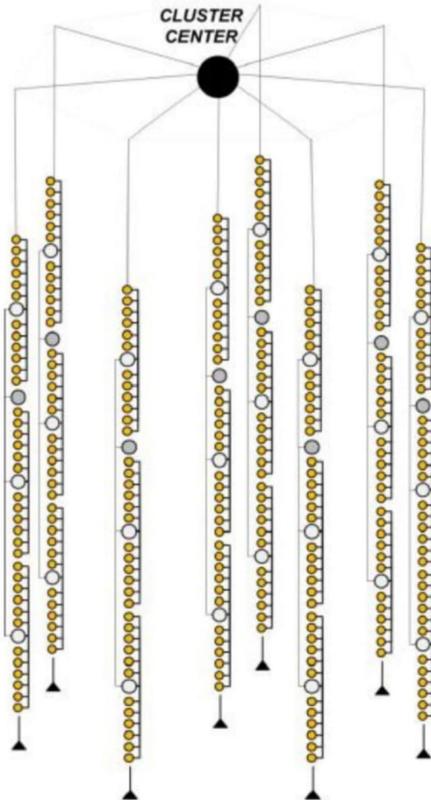
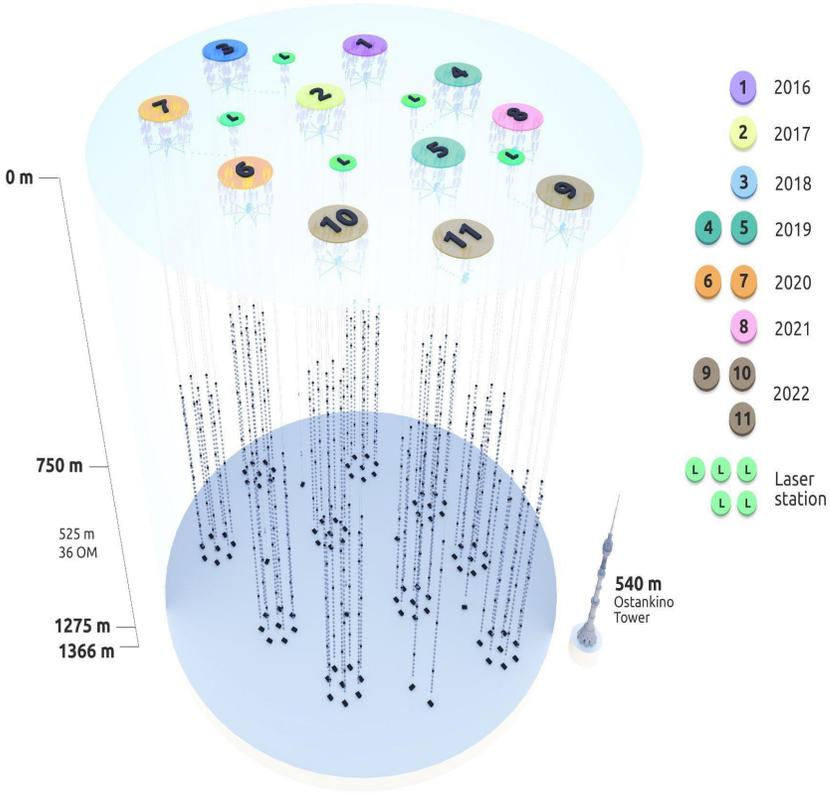
I. Baikal-GVD



Baikal-GVD



Детектор → 11 кластров → 8 струн → 36 оптических модулей (ОМ)



Живая калибровка положений ОМ (точность ~20 см)

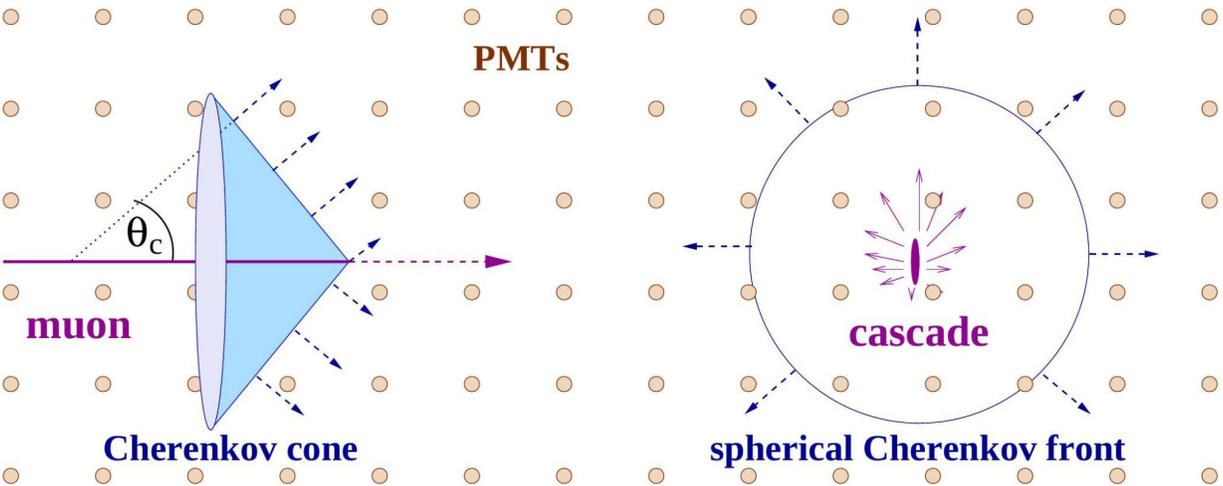
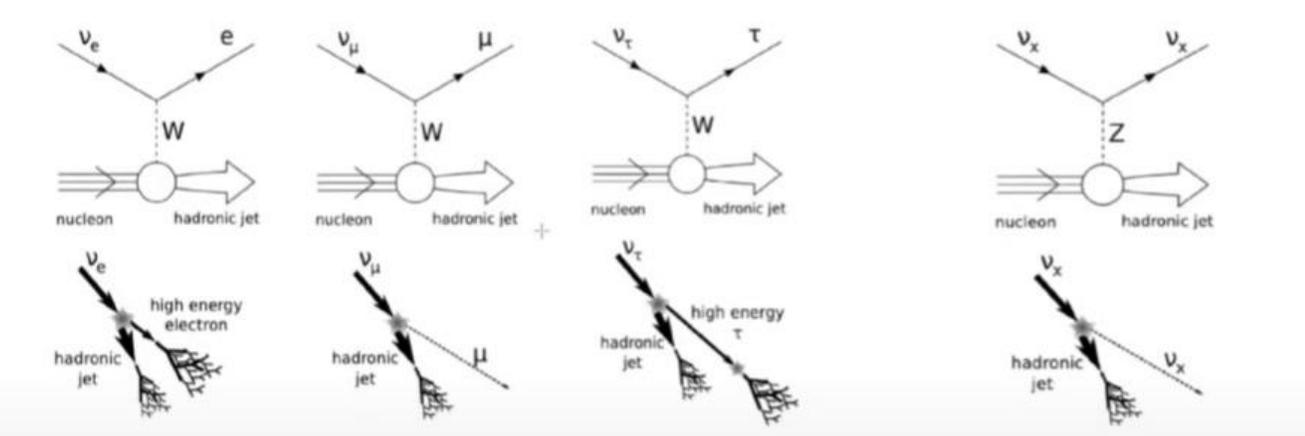
Точность определения времени срабатываний ~2 нс.

Baikal-GVD

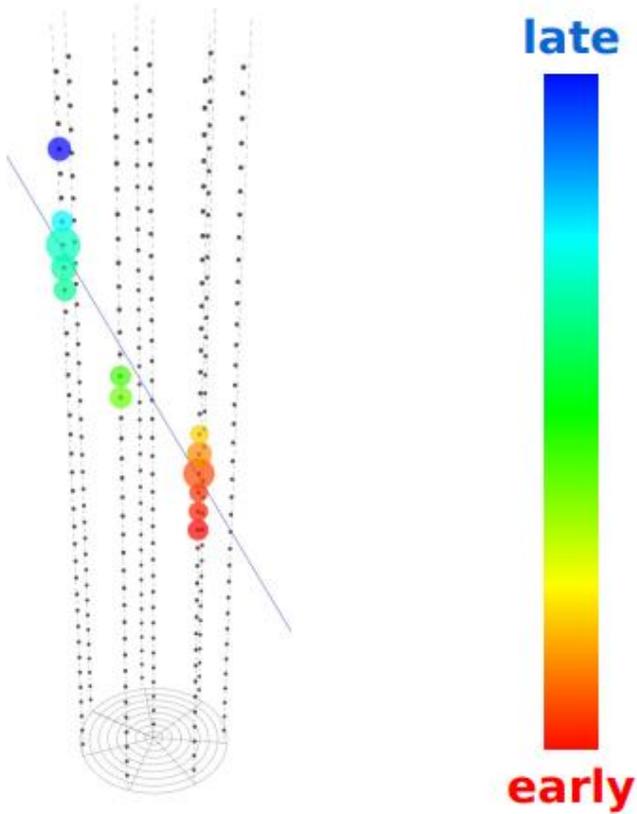
ОМ регистрируют фотоны от черенковского излучения (+ шум)

“Заряженные токи”: взаимодействие через W^\pm бозоны

“Нейтральный ток”: Взаимодействие через Z^0 -бозон



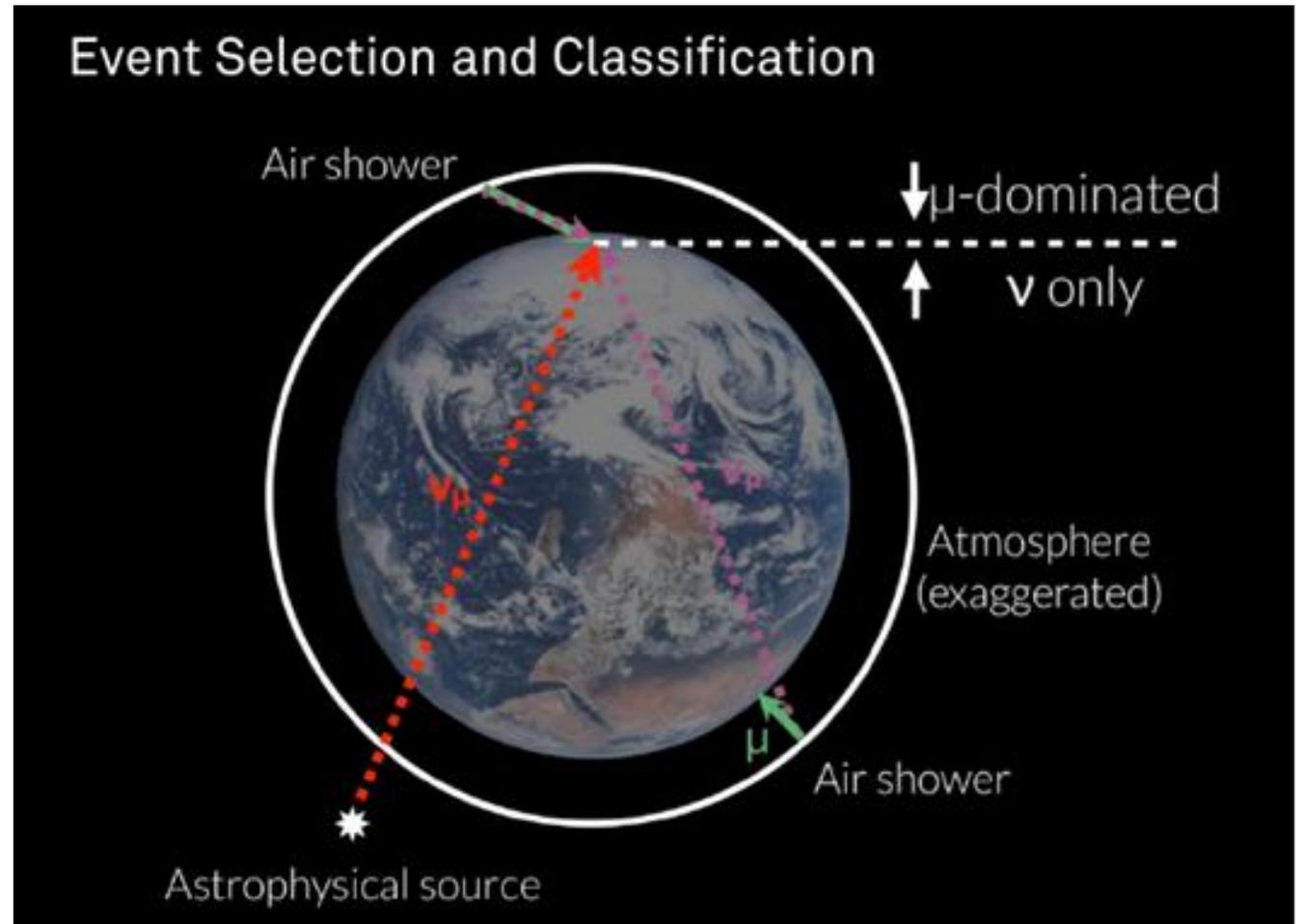
Пример трэкового события

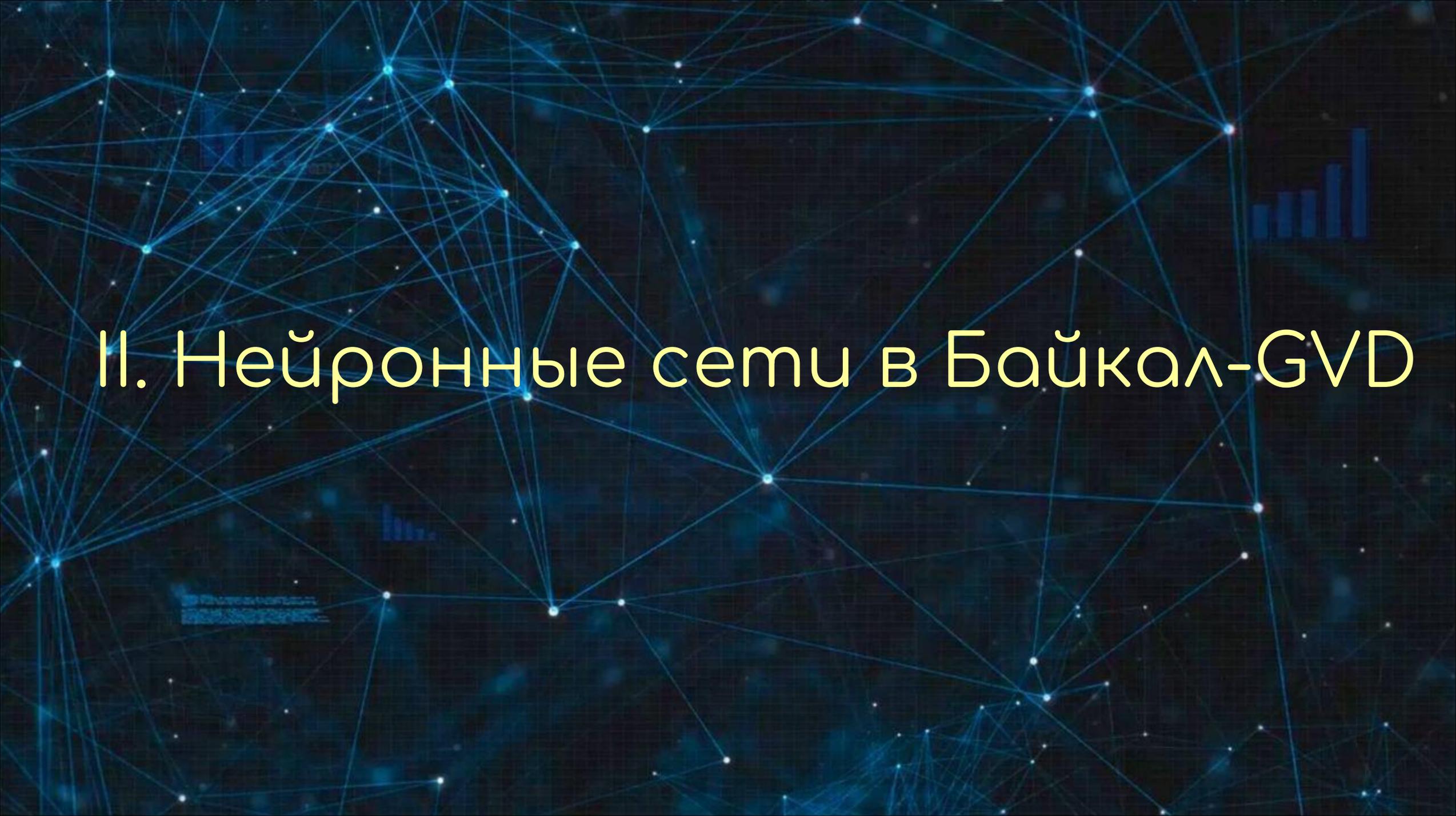


μ и ν индуцированные события

События:

- ШАЛ-индуцированные (атмосферные μ)
- нейтрино индуцированные (астрофизические или атмосферные нейтрино)





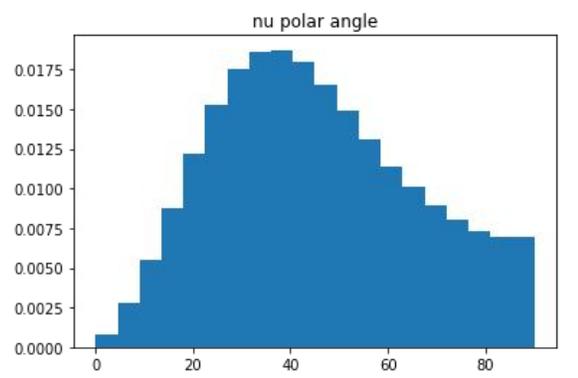
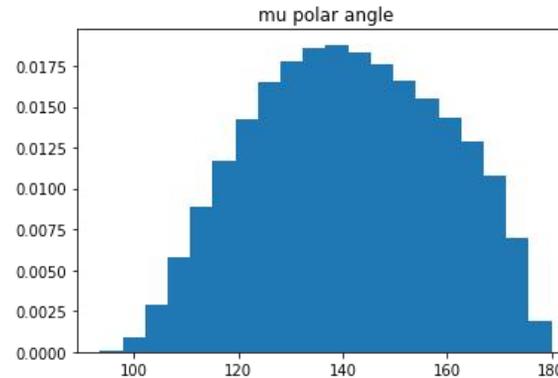
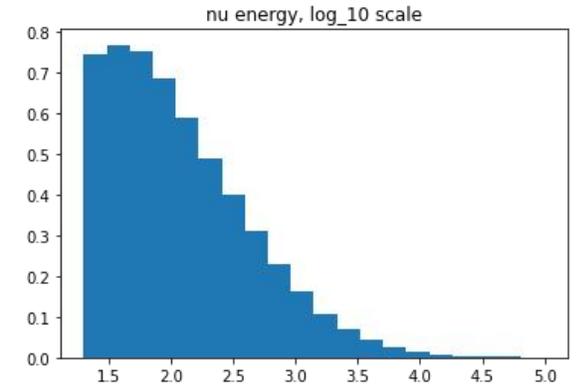
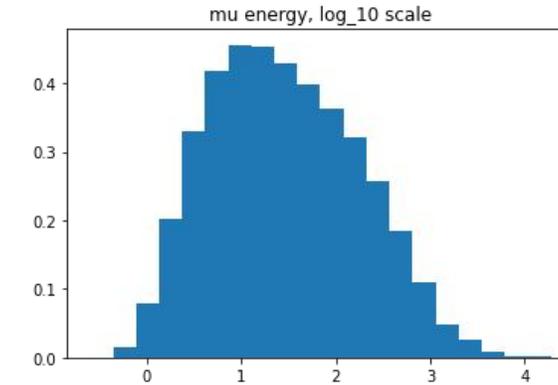
II. Нейронные сети в Байкал-GVD

Используемое Монте-Карло^[1]

Два типа событий:

- 1) Мюоны от космических лучей (μ)
- 2) ν_μ -индуцированные события

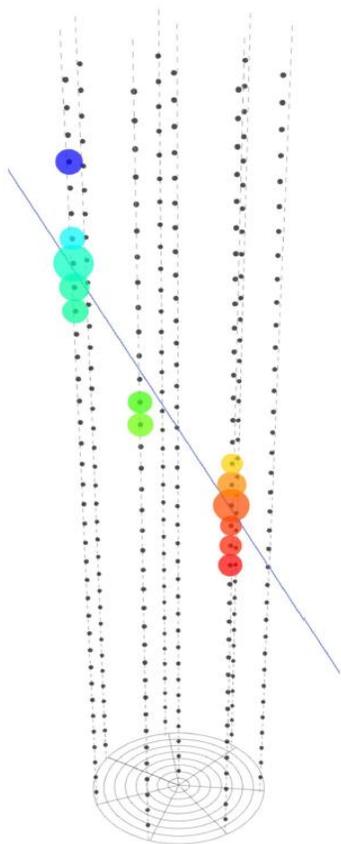
Полная симуляция развития ливней и распространение частиц в воде.



Для события известны: параметры первичной частицы (тип, энергия, угол).

Для хитов известны: координаты x , y , z ; время активации; интегральный сигнал; источник сигнала (какой μ) и его тип (трэк/каскад)

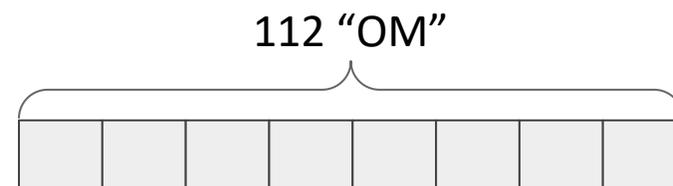
Представление данных



МК: один кластер, фиксированная геометрия



Причинное (временное)



Упорядочивание по времени

Добавление гауссова шума

1D свертки и рекуррентные сети

1. Подавление шумовых срабатываний ОМ

Мотивация:

Шумовые срабатывания составляют $\sim 85\%$ от собранных данных, уровень сигнала $\sim 1-3$ ф.э.

Кат на уровень сигнала слишком сильный: уберет сигнальные хиты, важные для реконструкции.

Нужно убирать как можно больше шума с наименьшей возможной ошибкой.

Два режима работы нейросетей:

А) “Полный”: пытаемся идентифицировать все сигнальные хиты

Б) “Трэковый”: ищем сигнальными хиты только вблизи трека мюона ($t_{res} < 20\text{ns}$)

1. Используемая нейросеть

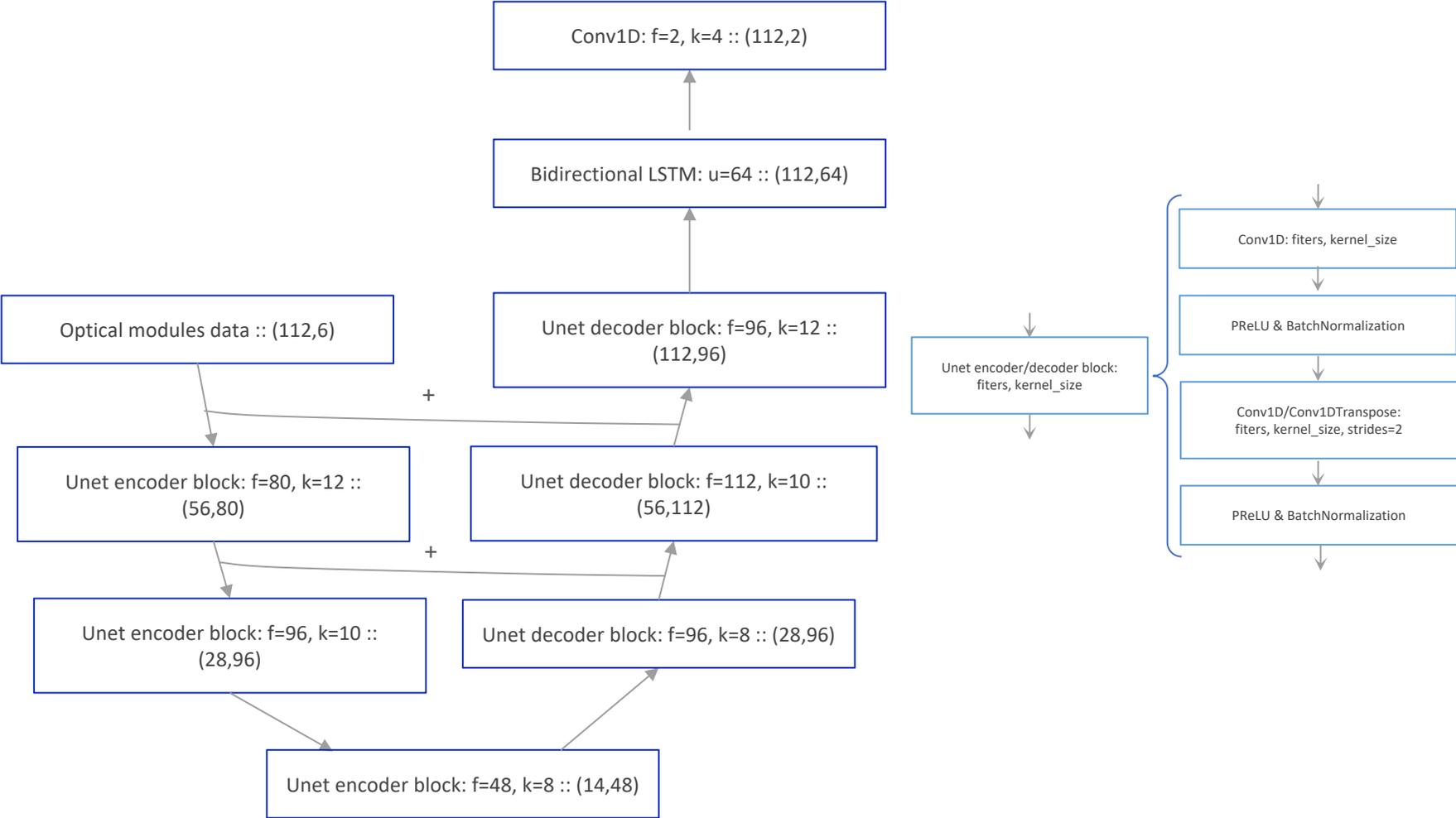
Входные параметры:

- 1-3) координаты x, y, z
- 4) интегральный сигнал
- 5) время активации
- 6) маска (вспомогательный/настоящий хит)

Выход:

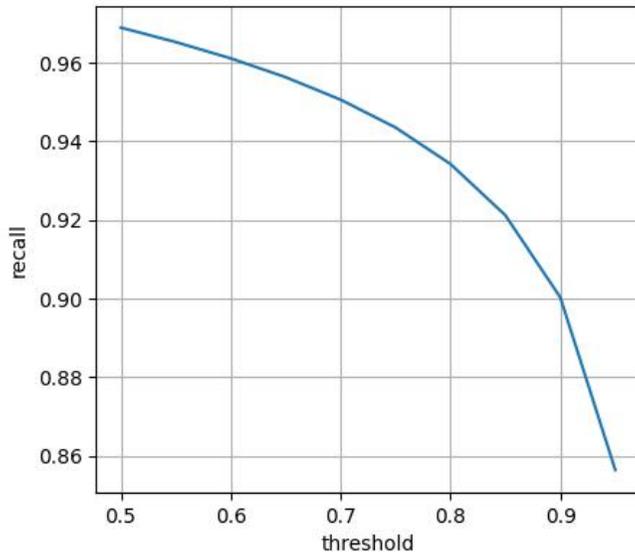
Для каждого ОМ, уверенность НС в том, что это сигнальный хит, $p \in [0;1]$

Unet: изучение локальных и глобальных свойств (горизонтальные связи и вертикальные связи)

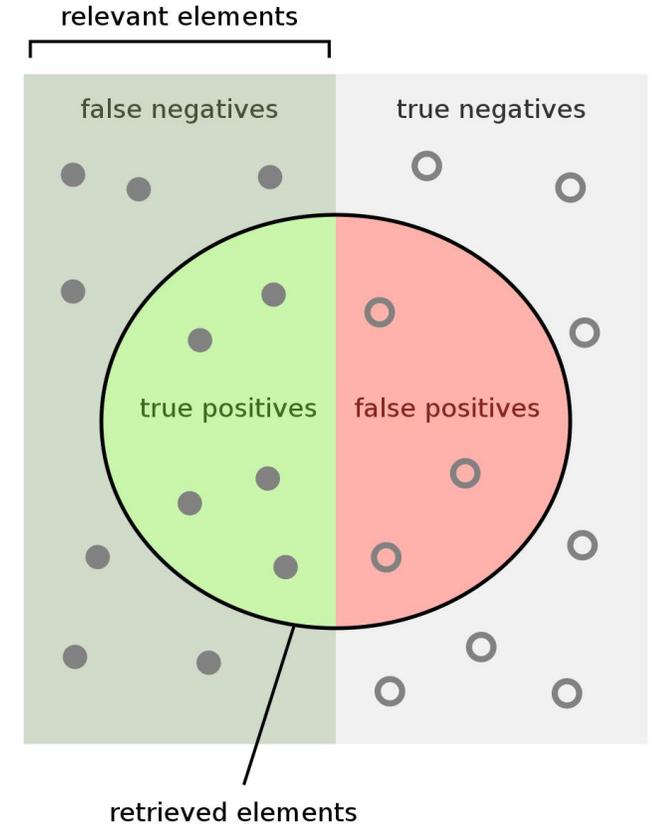
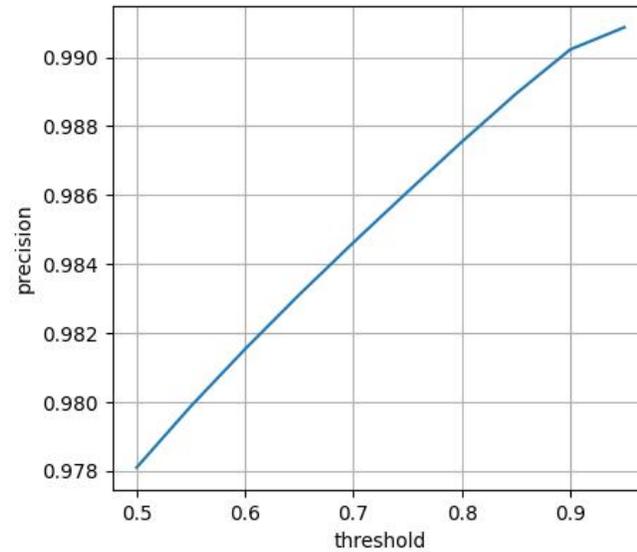


1. Метрики

Полнота



Точность



Точность (precision) и полнота (recall) **лучше на 4% и 1% соответственно!**

Скорость обработки быстрее до 10^3+ раз.

How many retrieved items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are retrieved?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

2. Разделение μ - и ν -индуцированных событий.

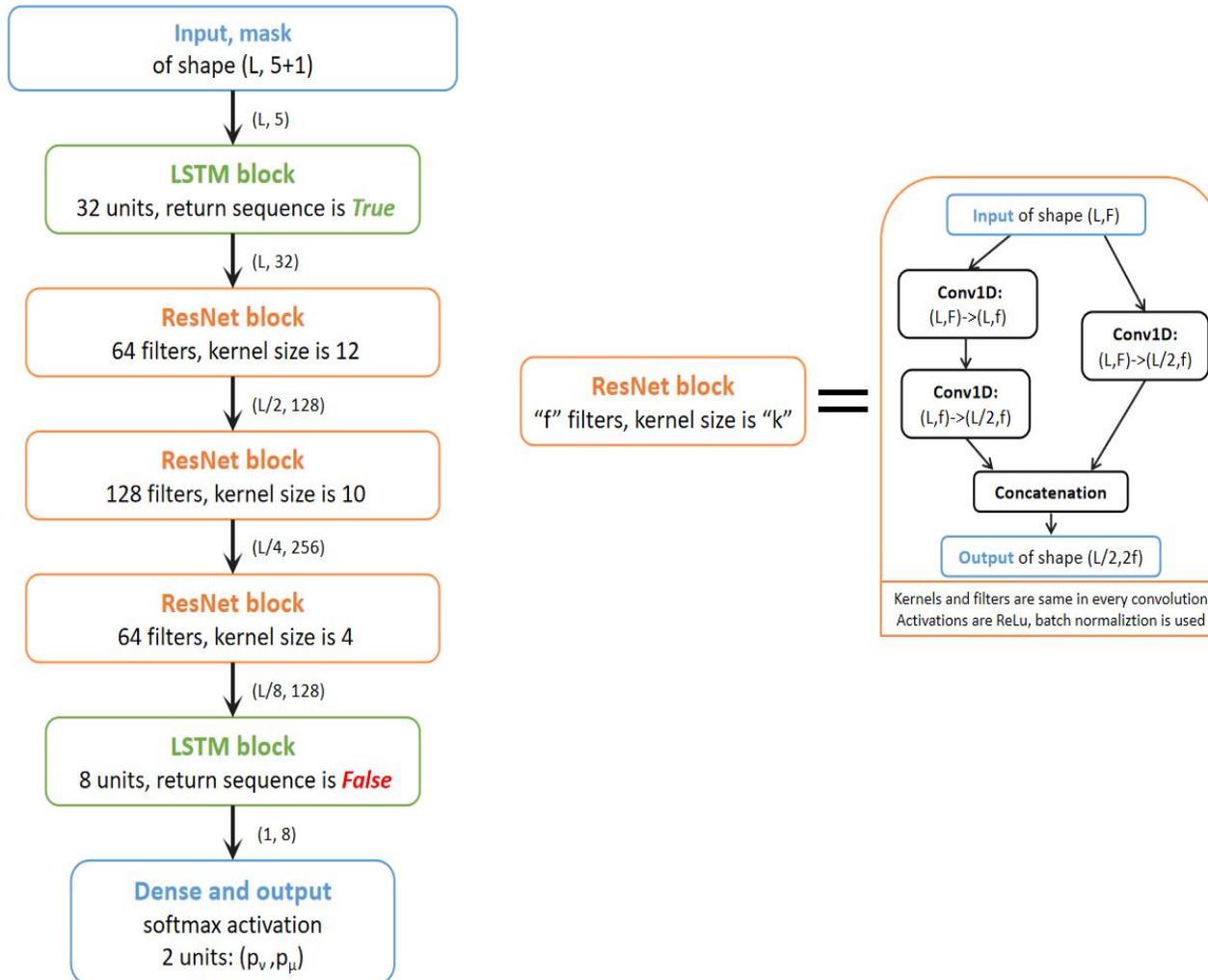
Мотивация:

Отношение μ к ν_μ событиям порядка 10^6 - 10^7 .

Стандартные каты достаточно жесткие.

Можно добиться лучшего разделения с более слабыми катами.

2. Используемая нейросеть



Функция потерь *focal loss*^[2]:

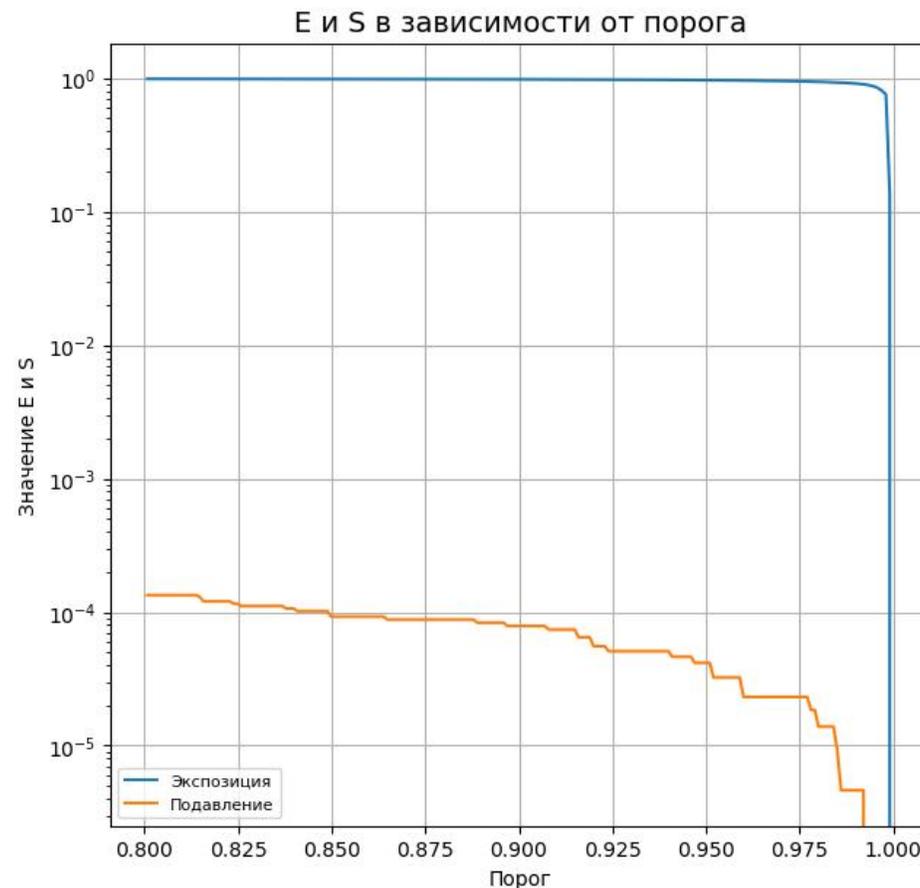
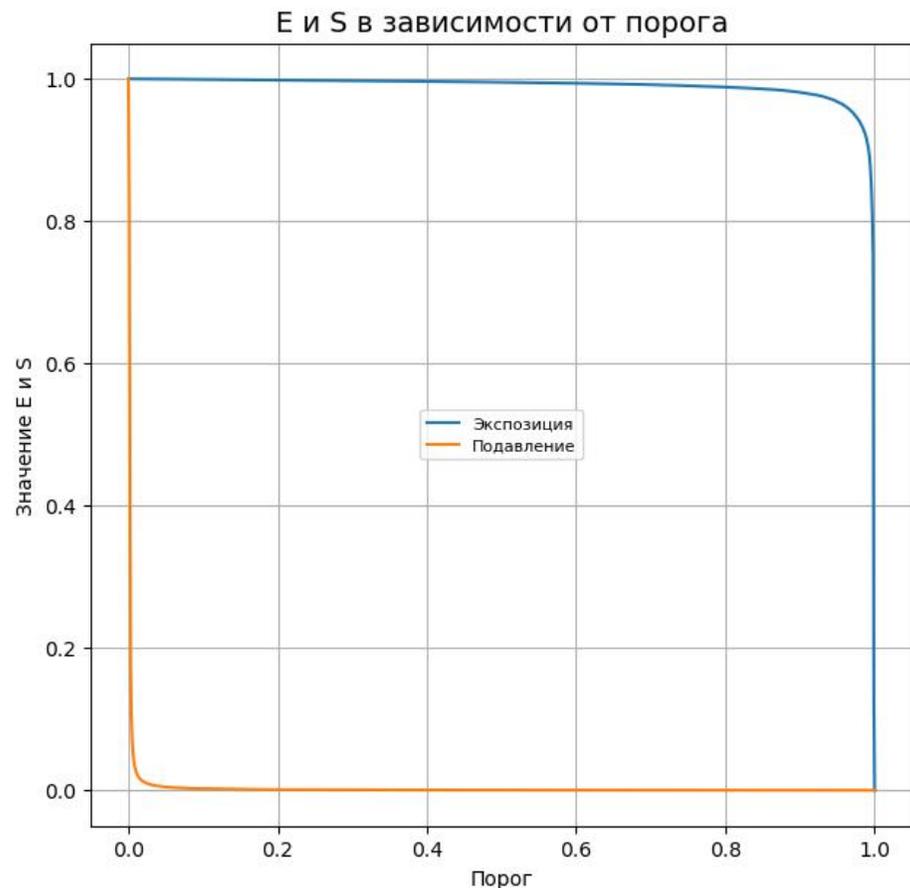
$$\text{Loss} = (1 - p_{\text{correct}})^2 * \text{bce}.$$

Гарантирует малое количество ложноположительных identификаций (ЛПИ).

2. Метрики

Экспозиция $E()$ - доля нейтрино справа от порога классификации

Подавление $S()$ - доля мюонов справа от порога классификации



Здесь $S = 0$ при пороге $0,993$. При этом $E = 90\%$.

2. Оценка потока нейтрино

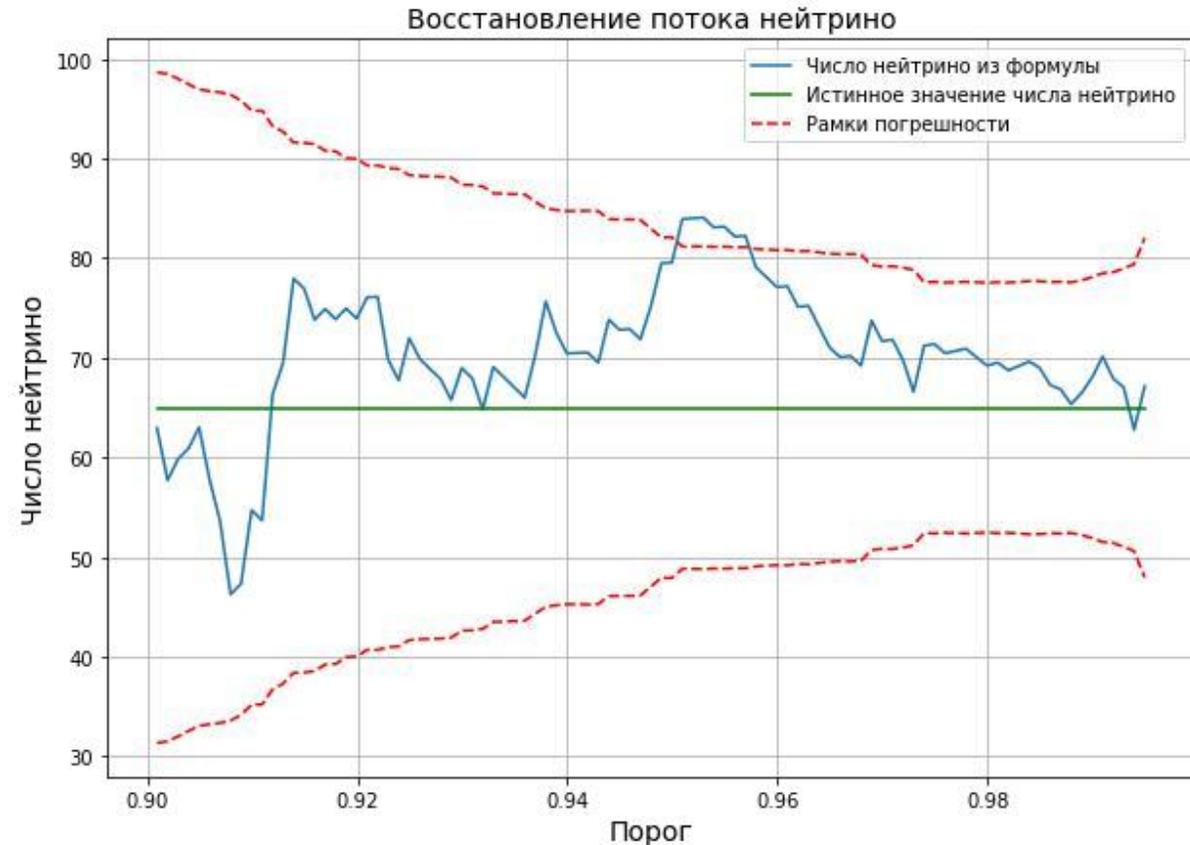
$$n_{\nu}^i \approx \frac{n^i(\xi) - S^0(\xi)n^i}{E^0(\xi) - S^0(\xi)}$$

S^0 , E^0 - подавление и экспозиция, оцененные на тестовом МК наборе данных.

ξ - порог классификации.

$n^i(\xi)$ - количество нейтринных событий, идентифицированных сетью правее порога.

Можно оценить ошибку!



Отношение μ к ν_{μ} : $\sim 500\,000$

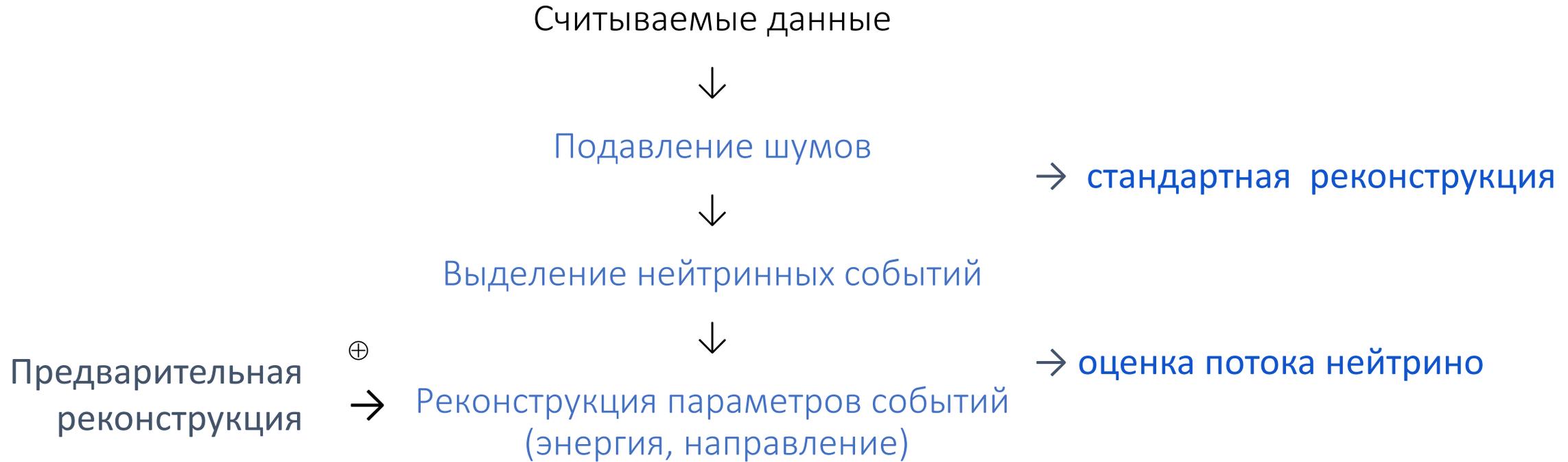
Заключение

- Нейронные сети позволяют улучшить качество реконструкции событий в эксперименте Байкал-GVD
- Планируется переход от МК к реальным экспериментальным данным
- Планируется разработка новых нейронных сетей для реконструкции энергии, выделения трековых и каскадных хитов, а также оптимизация существующих сетей

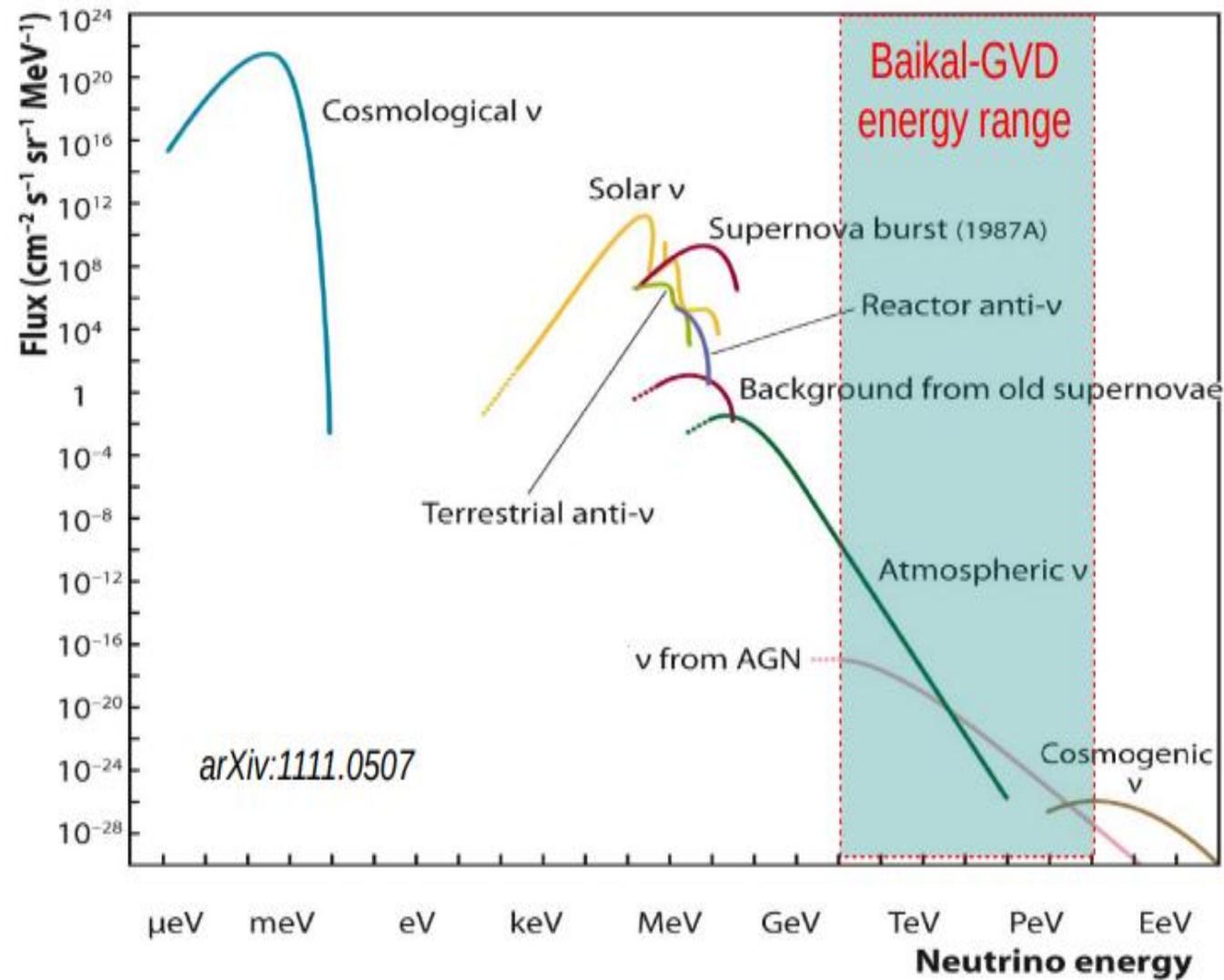
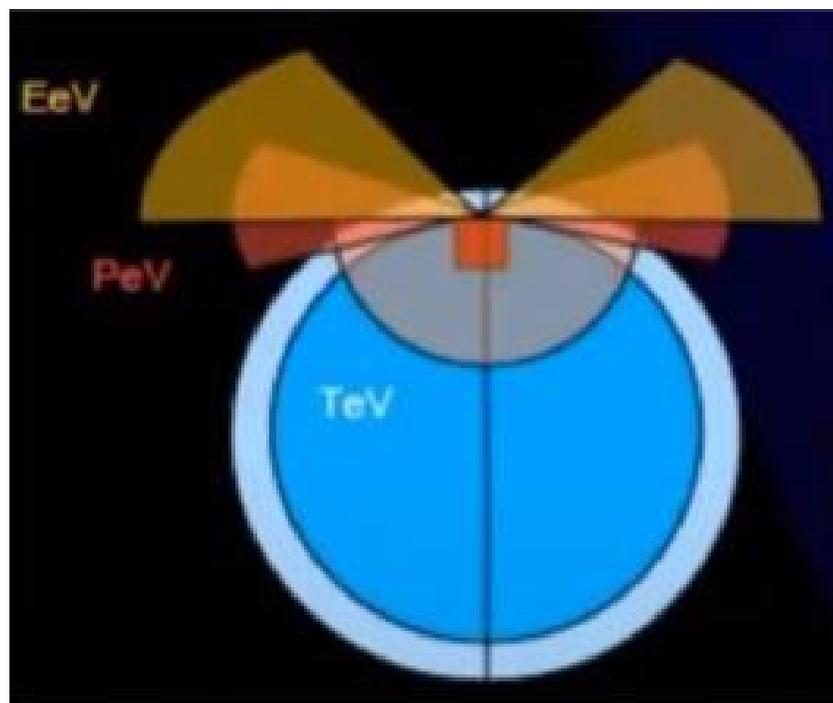
СПАСИБО ЗА ВНИМАНИЕ!

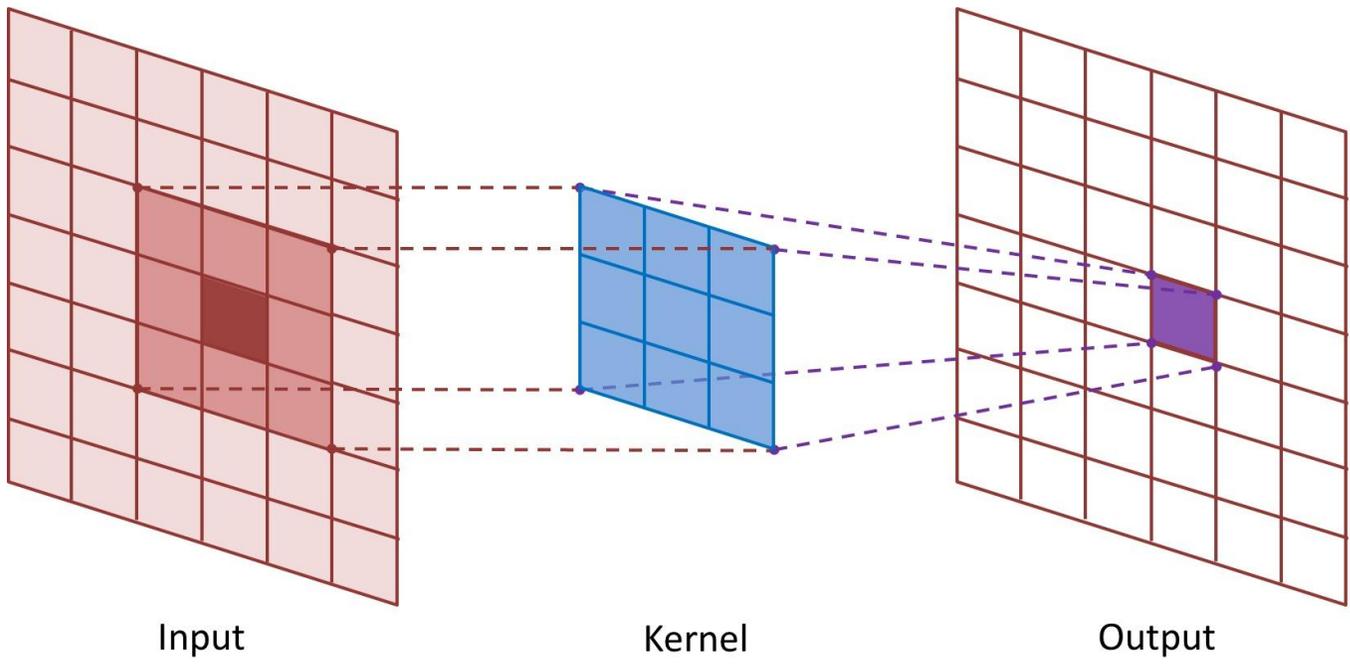
Backup

Общий план применения нейронных сетей

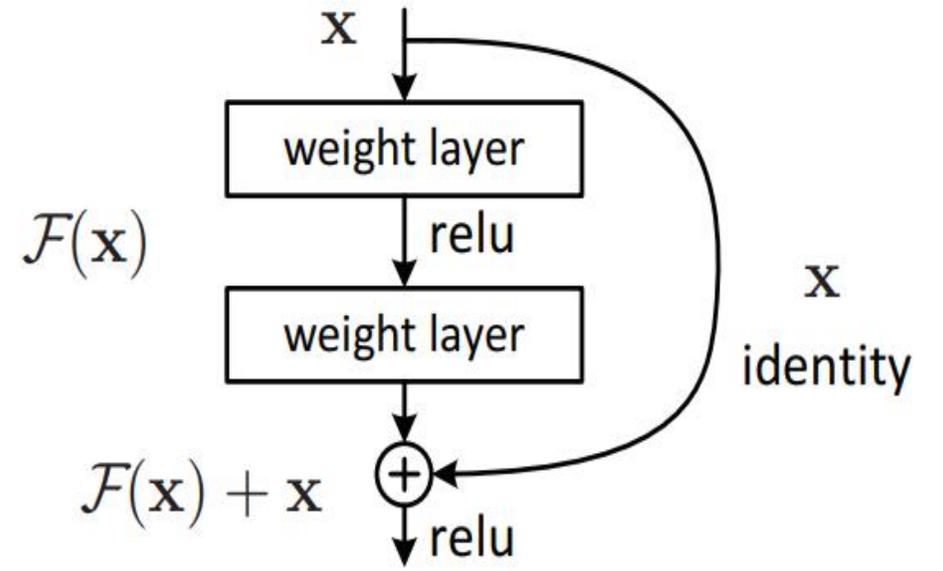


Классификация Нейтрино





R



Технические данные

$$\lambda_{scatterina}^{eff} \approx 480\text{м при } 475 \text{ nm} \quad \lambda_{absorption}^{max} \approx 24\text{м}$$

Трековые события :точность угла прилета $\approx 0, 25^\circ$
Каскадные события: разрешение $\approx 2^\circ$

Монте - Карло:

Взаимодействие нейтрино с ядрами : STEQ4M
(нейтрино с энергиями 10 ГэВ – 100ТэВ)

Прилет мюонов: программа CORSIKA 5.7 на модели
адронных вз-ий QGSJET

Распространение мюонов до Байкала : MUM v1.3u

Космические лучи: модель на базе KASCADE (240 ГэВ –
20 ПэВ)

Ошибка по времени 5 нс ; 30% по заряду

Focal loss

$$FL(p_t) = -\alpha_t (1-p_t)^\gamma \log(p_t)$$

Графики оценки потока

Исследование 1:

Набор	Число мюонных событий	Число нейтринных событий
Нулевой	20 171 379	262 995
Исследуемый	20 171 378	262 995

Табл.3. Конфигурация наборов для решения задачи восстановления потока.

Отношение мюон/нейтрино в исследуемом наборе равно 76.7.

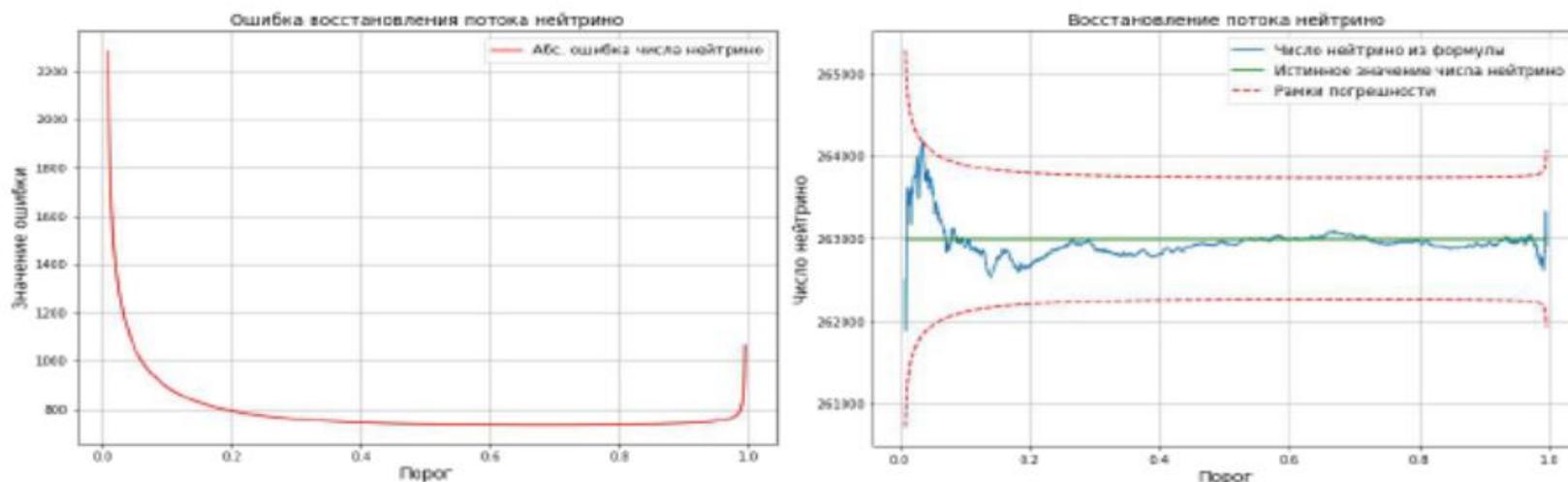


Рис. 23. Слева: $\sigma_{n_\nu^i}$ в зависимости от порога ξ . Справа: результат восстановления n_ν^i на исследуемом наборе и сравнение с истинным значением.

Минимальное значение ошибки достигается при $\xi = 0.674$.

Графики оценки потока

Исследование 2:

Набор	Число мюонных событий	Число нейтринных событий
Нулевой	13 447 586	473 392
Исследуемый	26 895 171	52 599

Табл. 4. Конфигурация наборов для решения задачи восстановления потока.

Отношение мюон/нейтрино в исследуемом наборе равно 511.3.

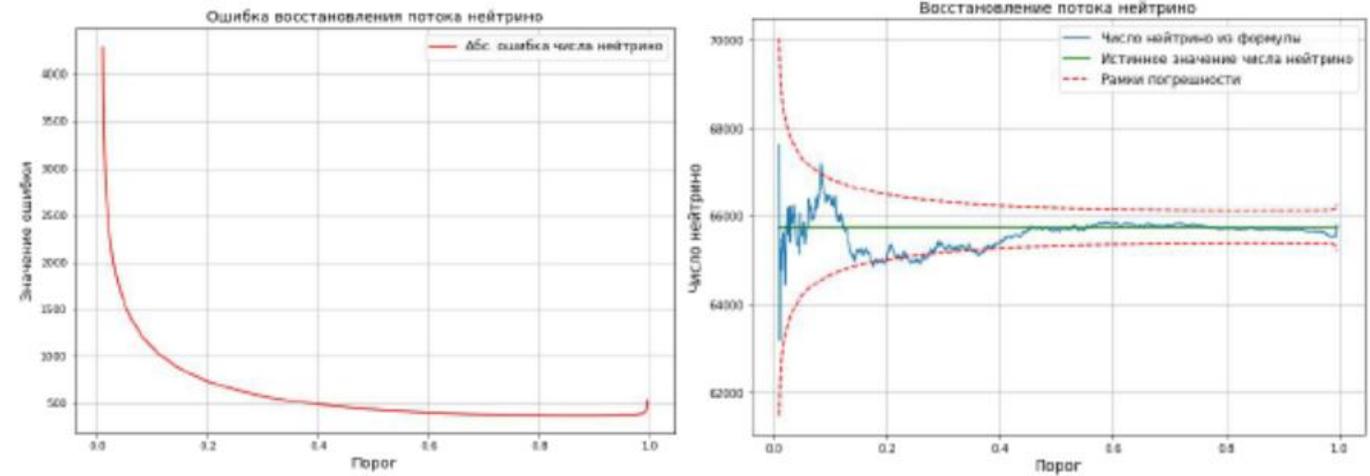


Рис. 24. Слева: $\sigma_{n_\nu^i}$ в зависимости от порога ξ . Справа: результат восстановления n_ν^i и сравнение с истинным значением.

Минимальное значение ошибки достигается при $\xi = 0.872$.

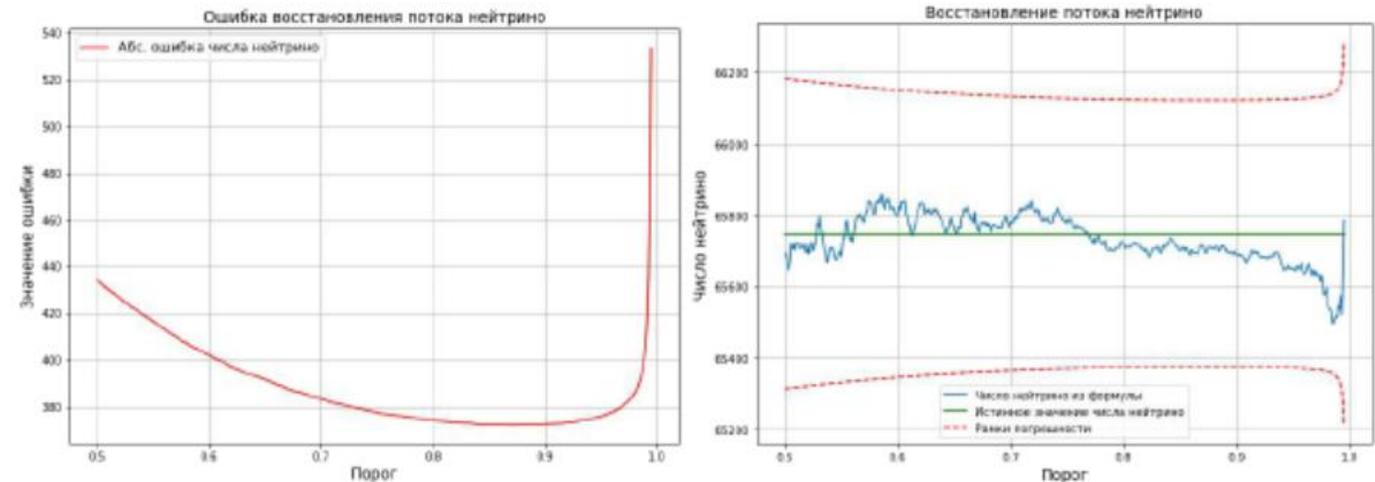


Рис. 25. Уточнение рисунка 24 при больших значениях ξ .

Графики оценки потока

Исследование 5:

Набор	Число мюонных событий	Число нейтринных событий
Нулевой	6 723 793	525 926
Исследуемый	33 618 964	65

Табл. 7. Конфигурация наборов для решения задачи восстановления потока.

Отношение мюон/нейтрино в исследуемом наборе равно 517 214.

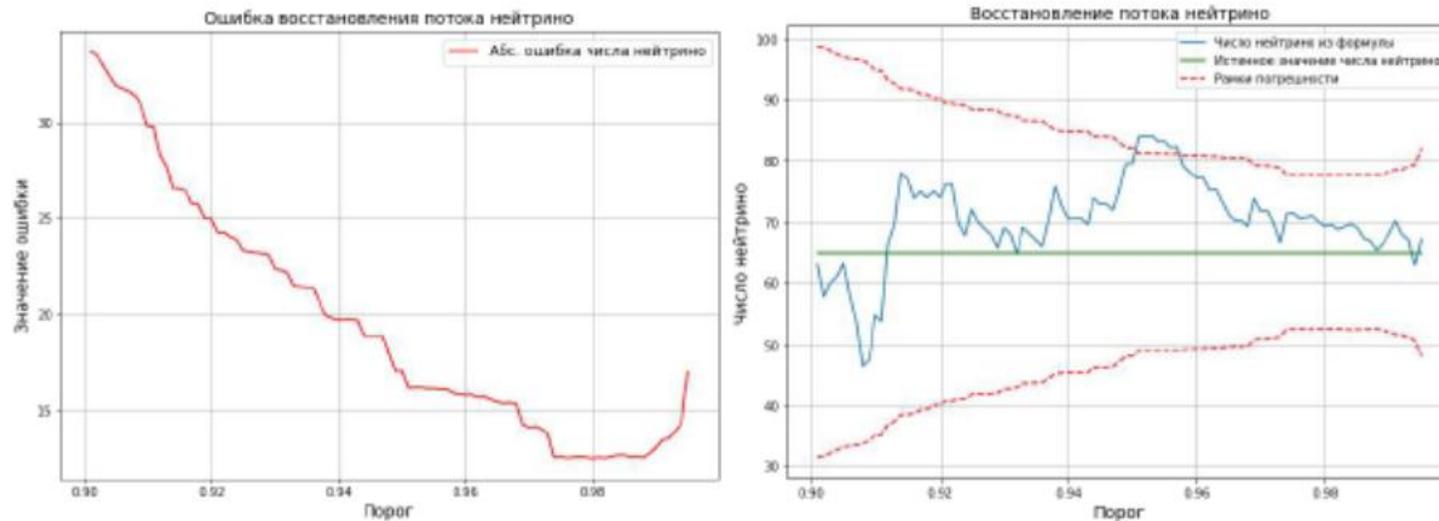


Рис. 29. Слева: $\sigma_{n_\nu^i}$ в зависимости от порога ξ . Справа: результат восстановления n_ν^i и сравнение с истинным значением.

Минимальное значение ошибки достигается при $\xi = 0.979$.

Вывод формулы ошибки

Для получения ошибки выражения (12) воспользуемся стандартной формулой оценки погрешности сложной функции $f(x, y)$ [16, стр. 171]:

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right)Cov_{x,y} \quad (23)$$

В нашем случае $f(x, y) = n_v^i(E^0(\xi), S^0(\xi), n^i(\xi), n^i)$, причём $E^0(\xi)$ и $S^0(\xi)$ - независимые случайные величины (события от нейтрино и от мюонов происходят независимо). Значит их ковариация равна нулю. Также равна нулю ковариация: $Cov(E^0(\xi), n^i(\xi))$, $Cov(S^0(\xi), n^i(\xi))$, $Cov(E^0(\xi), n^i)$, $Cov(S^0(\xi), n^i)$, т.к. на самом деле экспозиция и подавление, измеренные на нулевом наборе (например МК данные), и число частиц, измеренное на i -ом наборе (например экспериментальные данные), - независимые случайные величины. Ненулевой будет ковариация $n^i(\xi)$ и n^i :

$$Cov[n^i(\xi), n^i] = M[(n_v^i(\xi) + n_\mu^i(\xi))(n_v^i + n_\mu^i)] - M[n_v^i(\xi) + n_\mu^i(\xi)]M[n_v^i + n_\mu^i] \quad (24)$$

Как можно убедиться из распределения (14):

$$M[n_v^i(\xi)] = E(\xi)v; \quad M[n_\mu^i(\xi)] = S(\xi)\mu \quad (25)$$

$$M[n_v^i(\xi)n_\mu^i] = E(\xi)(v + v^2); \quad M[n_\mu^i(\xi)n_\mu^i] = S(\xi)(\mu + \mu^2) \quad (26)$$

Заодно посчитаем дисперсии:

$$D[n_v^i(\xi)] = E(\xi)v; \quad D[n_\mu^i(\xi)] = S(\xi)\mu; \quad D[n^i(\xi)] = E(\xi)v + S(\xi)\mu \equiv n(\xi) \quad (27)$$

Нейтрино и мюоны прилетают независимо, поэтому:

$$M[n_v^i(\xi)n_\mu^i] = 0; \quad M[n_\mu^i(\xi)n_v^i] = 0 \quad (28)$$

Раскрывая скобки в (24) и подставляя выражения (25), (26) и (28), получим:

$$Cov[n^i(\xi), n^i] = E(\xi)v + S(\xi)\mu \equiv n(\xi) \quad (29)$$

Теперь, по формуле (23) возьмём производные в выражении (12), подставим в качестве ошибок экспозиции и подавления их дисперсии (20) и (22) с индексом ноль (измеряем их на нулевом наборе), в качестве ошибок полного числа частиц - дисперсии (27) и ковариацию (29), измеренные на i -ом наборе. Получим:

$$\sigma_{n_v^i}^2 = \frac{(n^i(\xi) - S^0 n^i)^2}{(E^0 - S^0)^4} \cdot \frac{E^0(1 - E^0)}{n_v^0} + \frac{(n^i(\xi) - E^0 n^i)^2}{(E^0 - S^0)^4} \cdot \frac{S^0(1 - S^0)}{n_\mu^0} + \frac{n^i(\xi)}{(E^0 - S^0)^2} + \frac{(S^0)^2 n^i}{(E^0 - S^0)^2} - \frac{2S^0 n^i(\xi)}{(E^0 - S^0)^2} \quad (30)$$

Наша конечная цель состоит в поиске минимума для i -го набора ошибки (30) как функции порога ξ .