

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ МИФИ»  
(НИЯУ МИФИ)

УДК 539.121.667

ОТЧЁТ  
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

**Построение быстрого алгоритма фитирования гистограмм, близких к  
максвелловским функциям.**

Научный руководитель

к.ф.-м.н.,

\_\_\_\_\_ Д.С. Блау

Студент

\_\_\_\_\_ П. А. Бахтин

Москва 2021

# Содержание

Введение и физический смысл задачи	2
Постановка задач алгоритма	4
Описание предлагаемого алгоритма	6
Статистическое сравнение работоспособности и определения нужного параметра кодом	8
Сравнение времени работы алгоритмов	21
Выводы	22
Список используемых источников	23

# Введение и физический смысл задачи

Эксперимент ALICE - работает над описанием поведения кварк-глюонной плазмы, поиском стабильных сверхтяжёлых ядер трансурановых элементов. Одним из устройств, осуществляющих снятие экспериментальных данных является электромагнитный калориметр PHOS[1] - установка, регистрирующая падающие на них фотоны. Фотон с некоторой энергией попадает в детектор, где передаёт энергию электрону. На фотоумножителе, аналоговый сигнал усиливается многократно. Далее данный сигнал машинно оцифровывается, и наступает стадия первичной обработки: чтобы восстановить энергию и время прилёта фотона, необходимо вычислить параметры сигнала. Сам оцифрованный сигнал выглядит следующим образом:

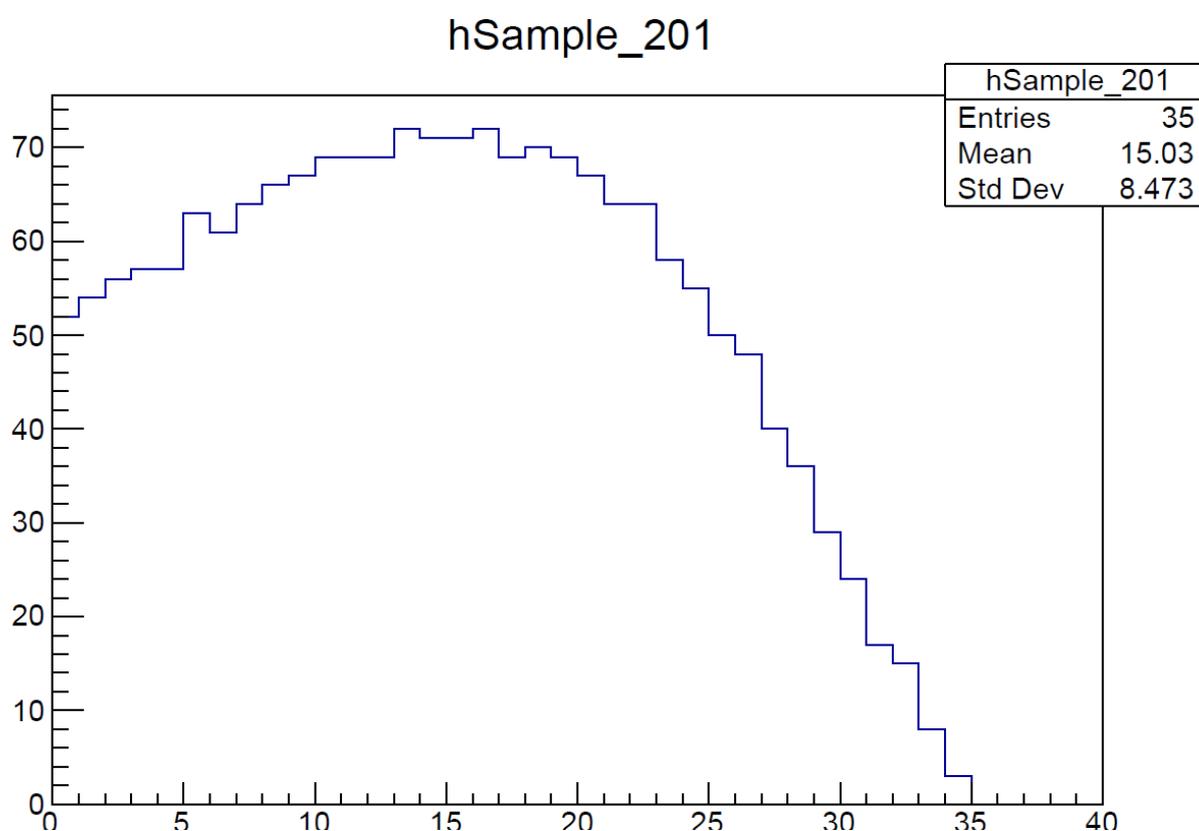


Рисунок 1 — Оцифрованный сигнал фотона снятый с калориметра

В работе [2] было показано, что в подавляющем большинстве слу-

чаев оцифрованный сигнал представляет собой вторую гамма-функцию, чья амплитуда связана с энергией фотона, а время прилёта указывает на геометрическое расположение события, позволяя оное идентифицировать.

Следовательно, основными искомыми параметрами фита являются: время  $\tau$  идентификации сигнала и амплитуда  $A$  сигнала. Дополнительным параметром могут выступать экспоненциальный множитель  $\beta$ , и, вообще говоря, степень  $t - \tau$ . Последняя принята равной двум, без потери точности

Следует отметить свойства сигнала: каждый бин - промежуток величиной 100 наносекунд (10 МГц АЦП), вертикальные значения лежат в пределах от 0 до 1024, каждая единица соответствует амплитуде, порождаемой фотоном с энергией 5 мэВ.

# Постановка задач алгоритма

Однако, поскольку для работы калориметра, то есть восстановления суммы потерянной на все фотоны энергии, необходима синхронная обработка указанных данных с их получением, к алгоритму фитирования возникают естественным образом некоторые требования, которые не позволяют использовать уже имеющееся средство программного пакета root для обработки.

Ранее, для определения искомых величин использовалось фитирование вершины с помощью параболы для поиска амплитуды, а затем фитирование прямой в окрестности точки начала.

Однако, поиск точки начала при помощи фитирования прямой в окрестности показал себя недостаточно быстрым и точным, и как следствие - возникла задача его заменить каким-либо другим способом, который будет соответствовать следующим требованиям:

- 1) Будет достаточно быстро (способен за 300 тактов процессора вычислить искомые параметры фита)
- 2) Разрешение параметра  $\delta\tau = 500$  пкс
- 3) Амплитудное разрешение  $\delta A = 0,5$  ADC

Подробнее рассмотрим связь легко обнаружимых точек функции и её основных параметров

$$f(t) = A \cdot (t - \tau)^2 \cdot \exp^{-\beta \cdot t} \quad (1)$$

Тогда время начала регистрации сигнала будет определяться выражением:

$$t = \tau \quad (2)$$

Выячив производную по времени, можно получить, что время достижения вершины однозначно связано с экспоненциальным множителем:

$$\beta = \frac{2}{t_{max} - \tau} \quad (3)$$

Соответственно "Опорное" для предлагаемого далее алгоритма значение можно вычислить, исходя из значения амплитуды согласно точному

значению максимального значения сигнала, и координаты оного.

# Описание предлагаемого алгоритма

Предложено следующее решение. Сначала, путём сравнения находится бин, представляющий собой максимум в гистограмме. Берётся его окрестность в виде 5 бинов справа и слева от него, после чего вершина фитируется параболой (полиномом 2 степени) методом  $\chi^2$ .

Таким образом, находится вершина функции, и получается соотношение, позволяющее однозначно установить все соотношения.

Следующей операцией, является минимизация по единственному неизвестному независимому параметру - величине  $\tau$ . Поскольку требуется достичь заранее нужной точности, алгоритм сделан следующим образом. Сначала ищется минимум функции по крупной сетке, с шагом в 0,5, затем этот шаг двукратно увеличивается, идёт уточнение значение минимума, после чего операция повторяется с увеличенной точностью. Причём, поскольку ось минимизации единственная, то можно искать именно устойчивый минимум, то есть, чтобы при движении и влево, и вправо величина функционала расстояния не увеличивалась. Тем самым, обходится процедура минимизации ФМП и сводится к одномерной минимизации функционала.

Код реализован в парадигме ООП, т.е. к алгоритму статистического сравнения точности фитирования подключен ф функция, данное фитирование ведётся. Примеры работы фитирования гистограмм:

Example Hystogram 207

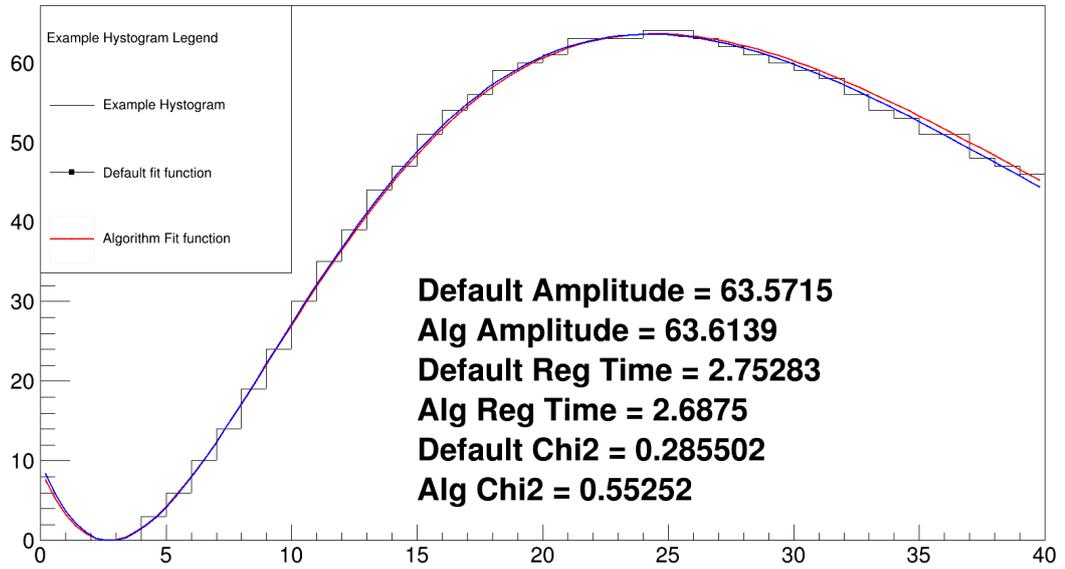


Рисунок 2

Example Hystogram 505

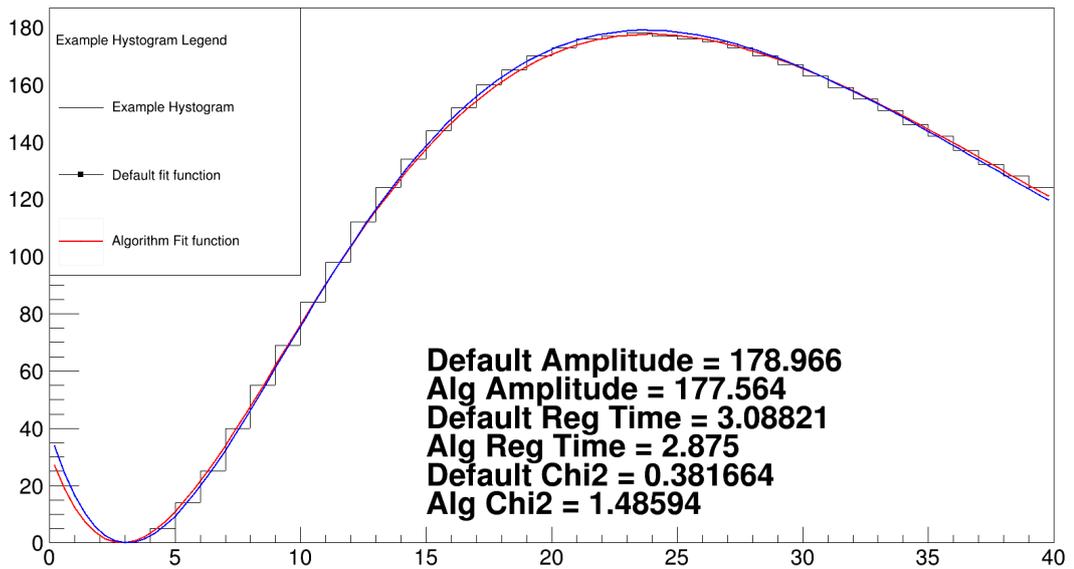


Рисунок 3

# Статистическое сравнение работоспособности и определения нужного параметра кодом

Соответственно, благодаря реализации кода в парадигме ООП, стало возможно провести анализ работоспособности алгоритма, его стабильности и точности, на основе 712 гистограмм оцифрованных сигналов.

Получены гистограммы значений параметров фитирования и критерия  $\chi^2$  к минимизации которого стремится алгоритм:

Pre-exp coef by every I histogram

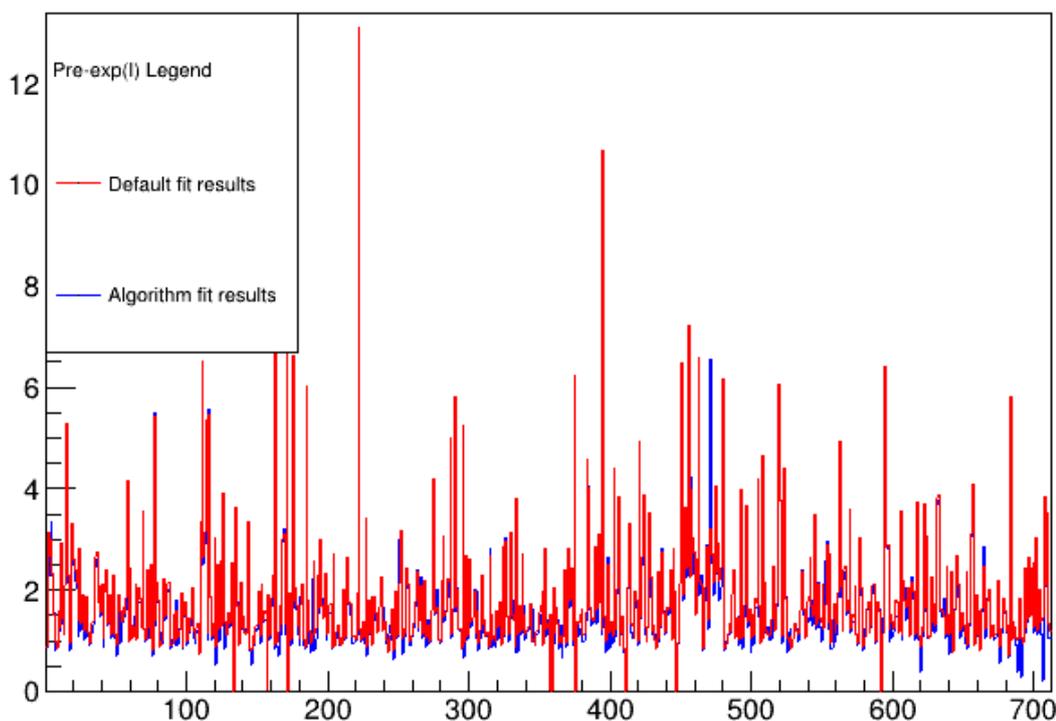


Рисунок 4 — Гистограмма с предэкспоненциальными коэффициентами сигналов

На их основе, получены границы для гистограмм, в которых можно получить статистику вычисления параметров сигналов двумя разными методами. Она выглядит следующим образом:

## Time of signal registration by every I histogram

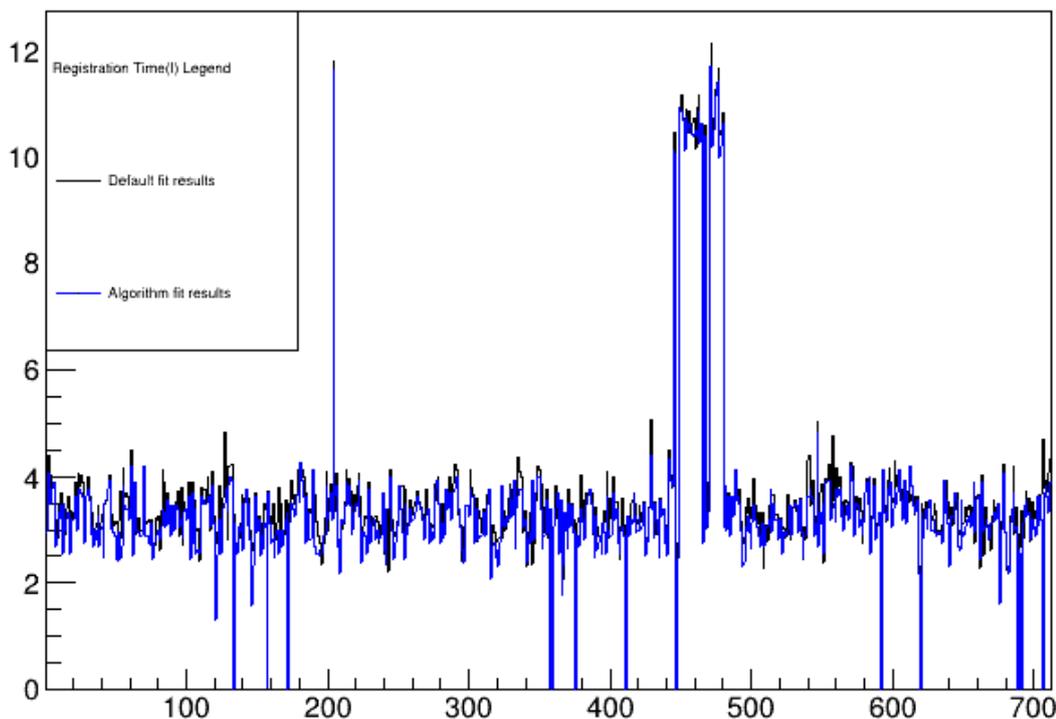


Рисунок 5 — Гистограмма с временами регистрации сигналов

Таким образом, можно оценить, насколько велики разбросы параметров сигналов. В частности, можно заметить, что вычисленные любыми методами коэффициенты степени экспоненты равны  $0.09^{+0.01}_{-0.01}$ . К сожалению, принять здесь как среднее значение параметра ничего не удастся, поскольку порождаемая относительная ошибка достаточно существенна

Следующим этапом проверки работоспособности предложенного алгоритма, следует составить гистограммы разностей по 3 параметрам. Используем комплексный параметр амплитуды сигнала для построения такого распределения, наряду с экспоненциальным коэффициентом и временем регистрации сигнала. Средняя разность и среднеквадратичное отклонение от оной будут представлены на рисунках. Также приведём эту же самую величину для параметра  $\chi^2/ndf$ .

Тем самым, можно говорить о том, что величины параметров, получаемых разными методами, имеют стабильное отклонение и разброс. Для выяснения природы первого, решено в дальнейшем использовать тестовую статистику. Второе же, разброс параметров, решено оценить при различ-

exp coef by every I histogram

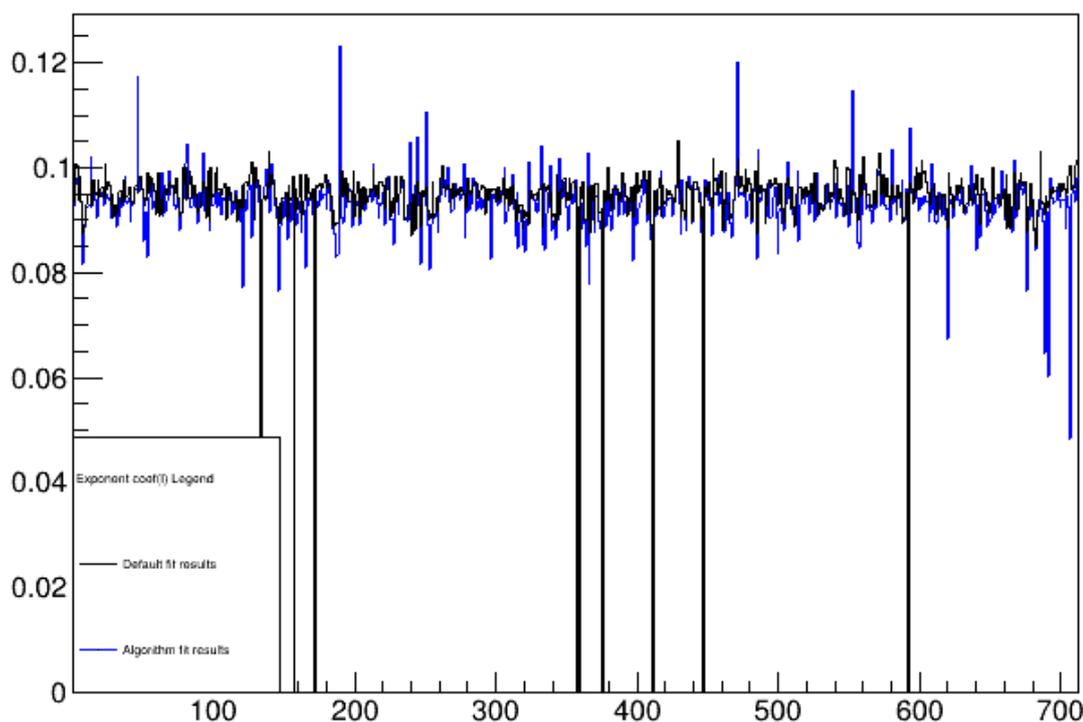


Рисунок 6 — Гистограмма с экспоненциальными показателями убывания сигналов

ных энергиях. Чтобы подтвердить случайную природу отклонений, построены гистограммы распределения по величине отклонения от амплитуды сигнала, соответствующей энергии.

Вторая представляет особый интерес, поскольку позволяет установить, как выглядит временное разрешение созданным методом в зависимости от энергии.

Полученное значение временного разрешения при энергии  $E = 0.5\text{GeV}$  составляет  $1.5\text{ns}$ . Такая величина соответствует тем результатам точности, которые удавалось получать ранее при регистрации фотона.

### Chi2 by every I histogram

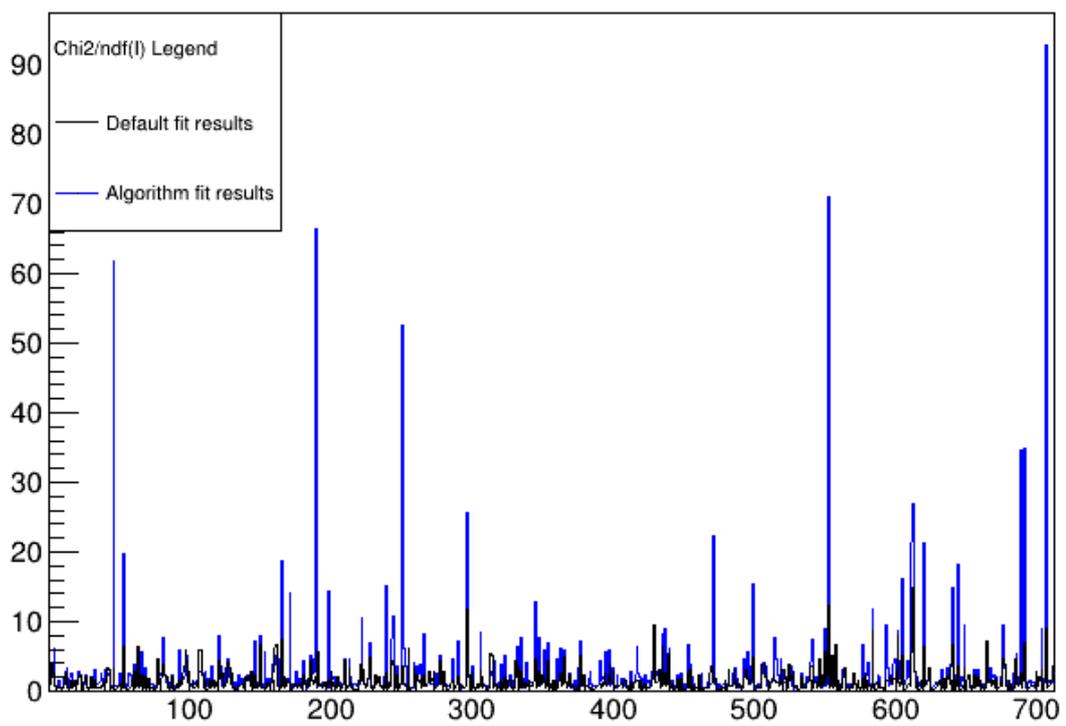


Рисунок 7 — Гистограмма с  $\chi^2$  результатов фитирования, Черная линия - menuite, синяя - авторский алгоритм

### Amplitude by every I histogram

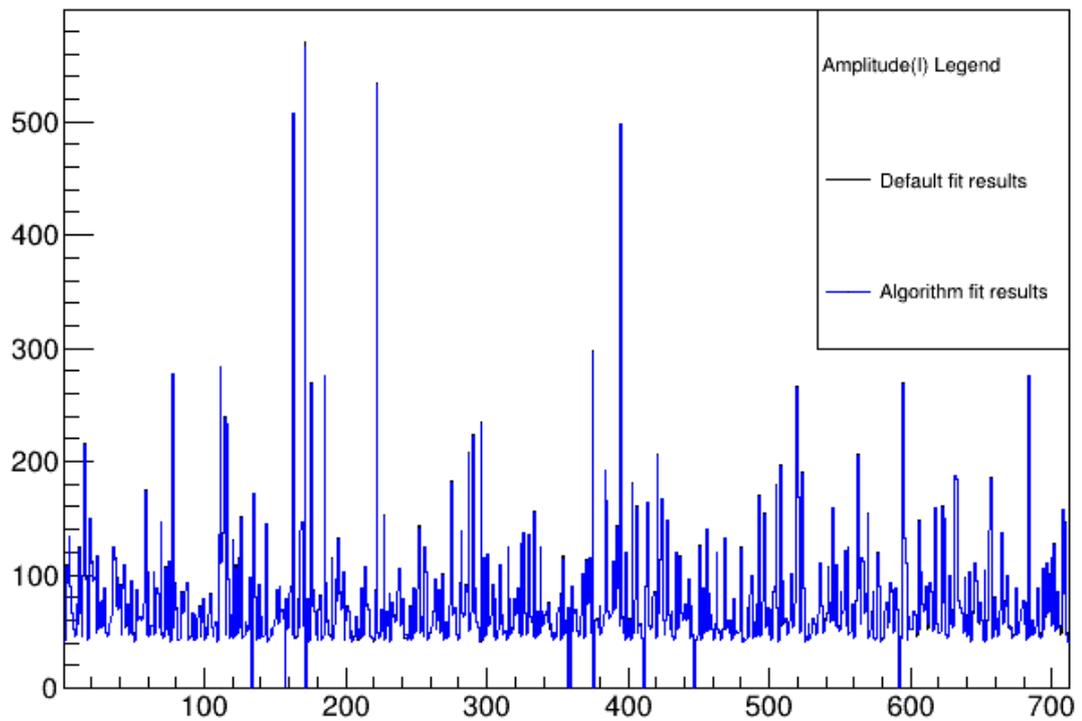


Рисунок 8 — Гистограмма с Амплитудами зарегистрированных сигналов

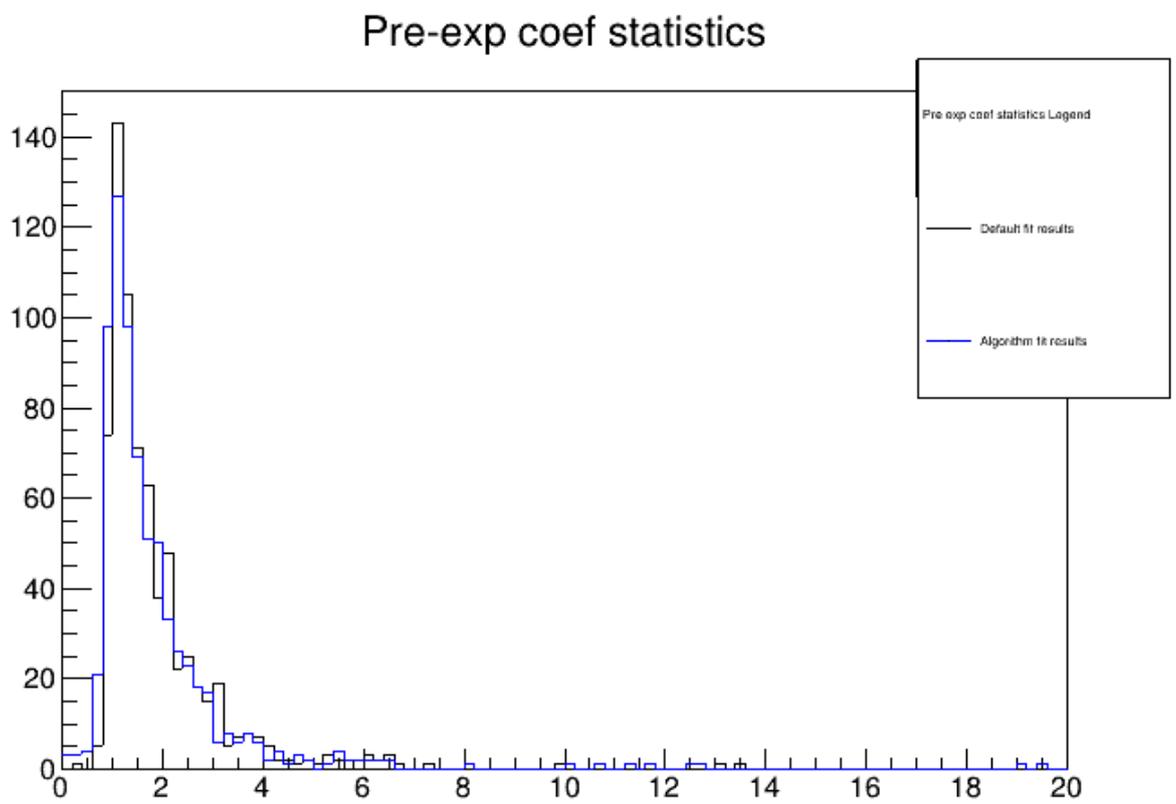


Рисунок 9 — Статистика распределения предэкспоненциального коэффициента

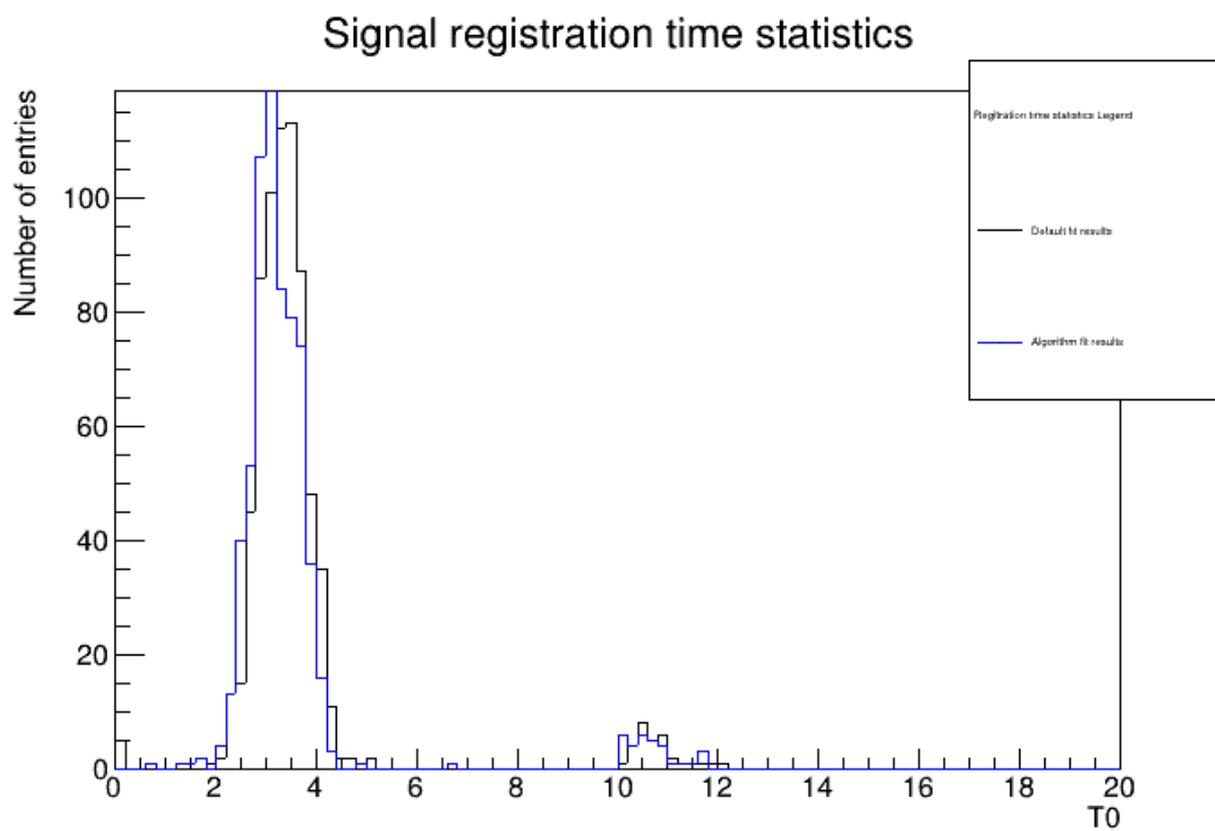


Рисунок 10 — Статистика распределения времени регистрации частицы

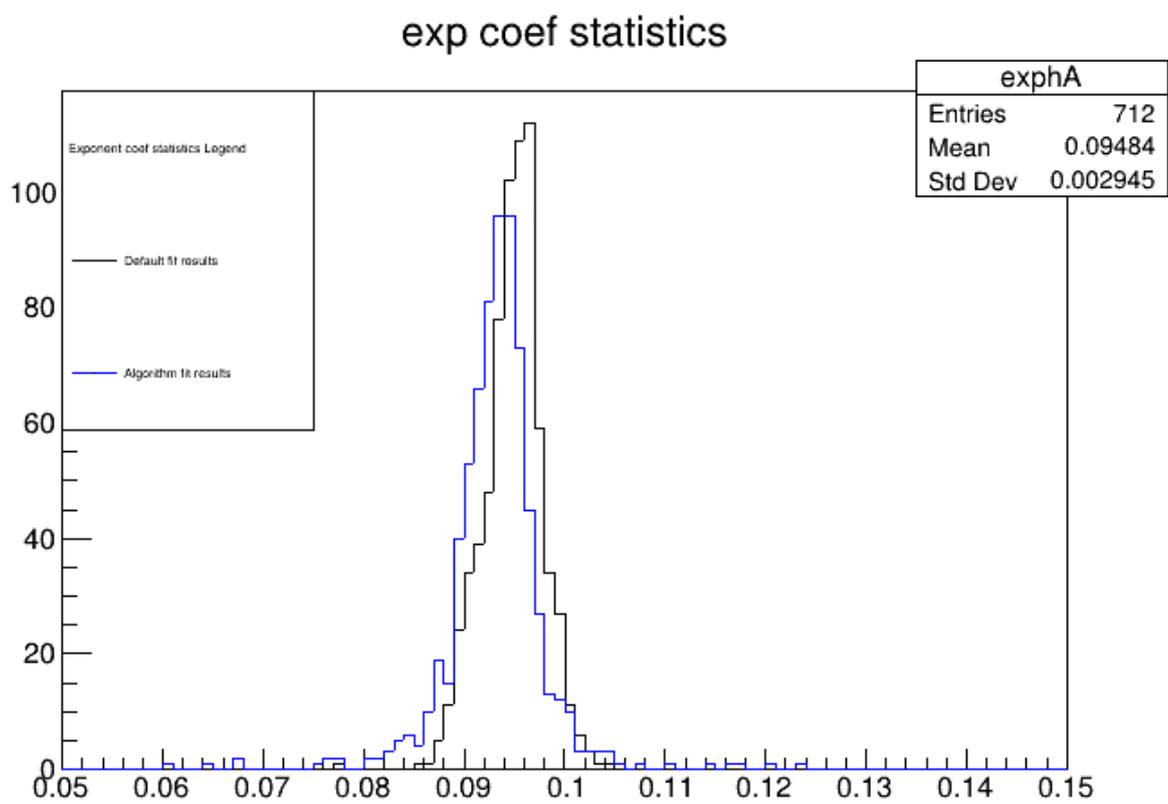


Рисунок 11 — Статистика распределения экспоненциального коэффициента

### Chi2 statistics

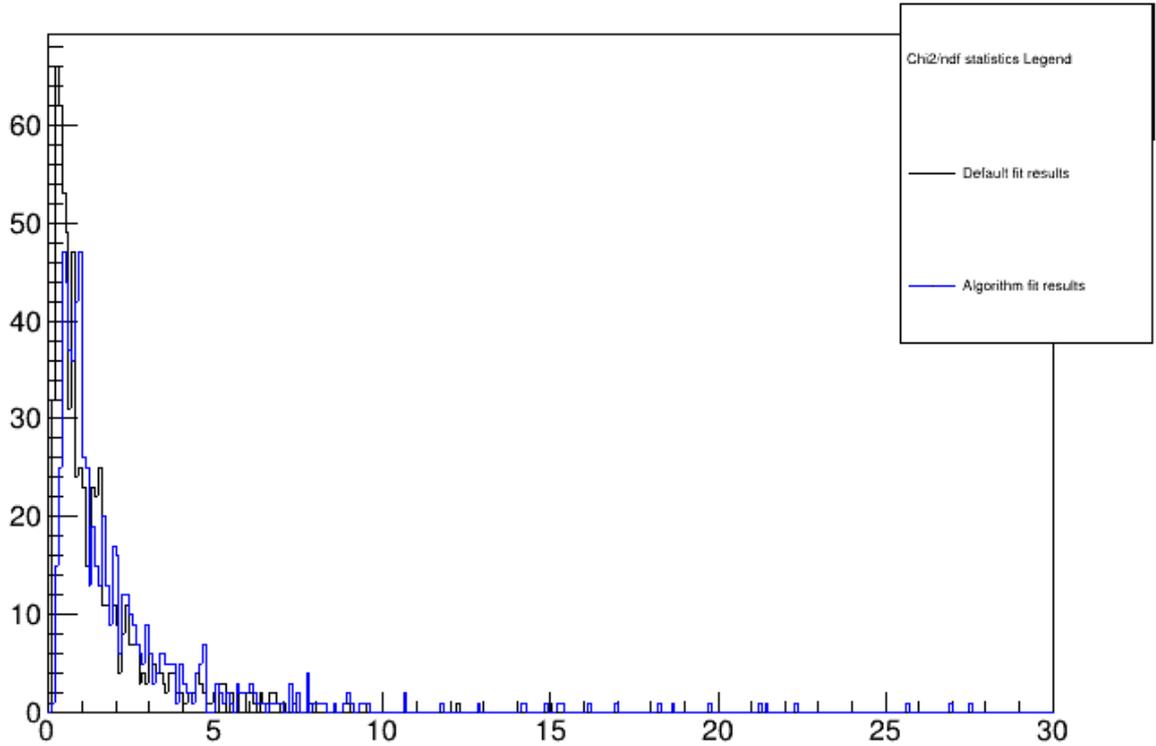


Рисунок 12 — Статистика распределения величины  $\chi^2$

### Amplitude statistics

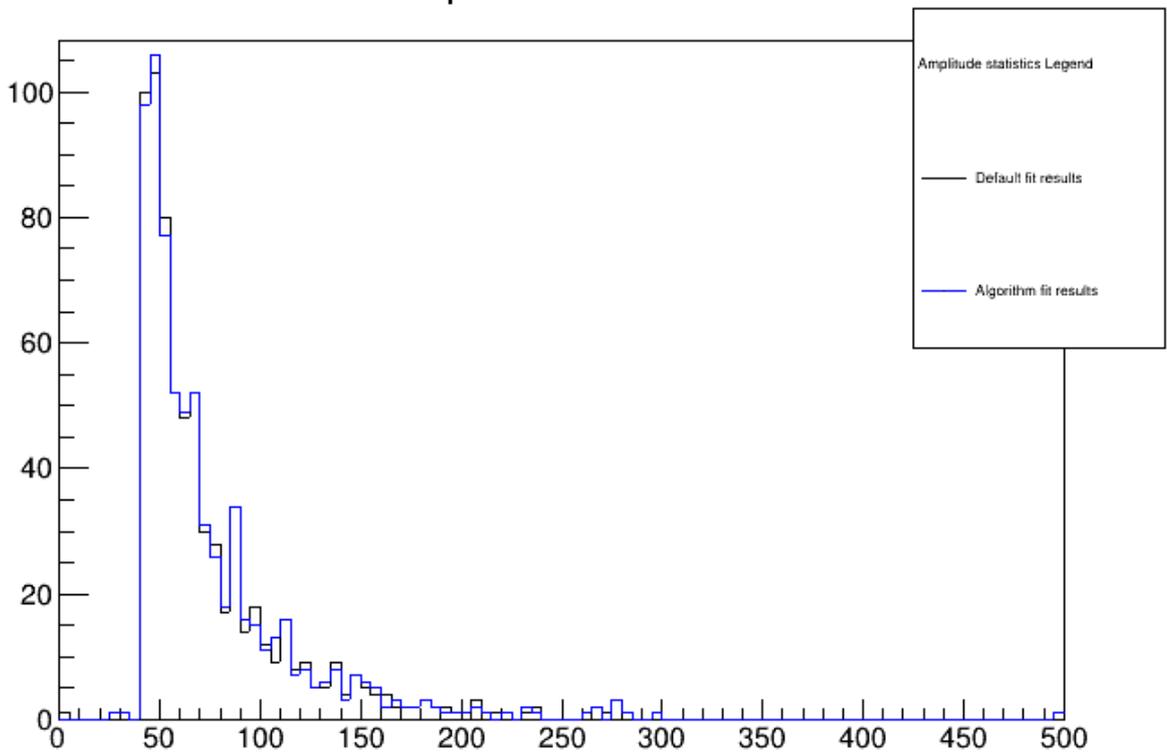


Рисунок 13 — Статистика распределения амплитуд

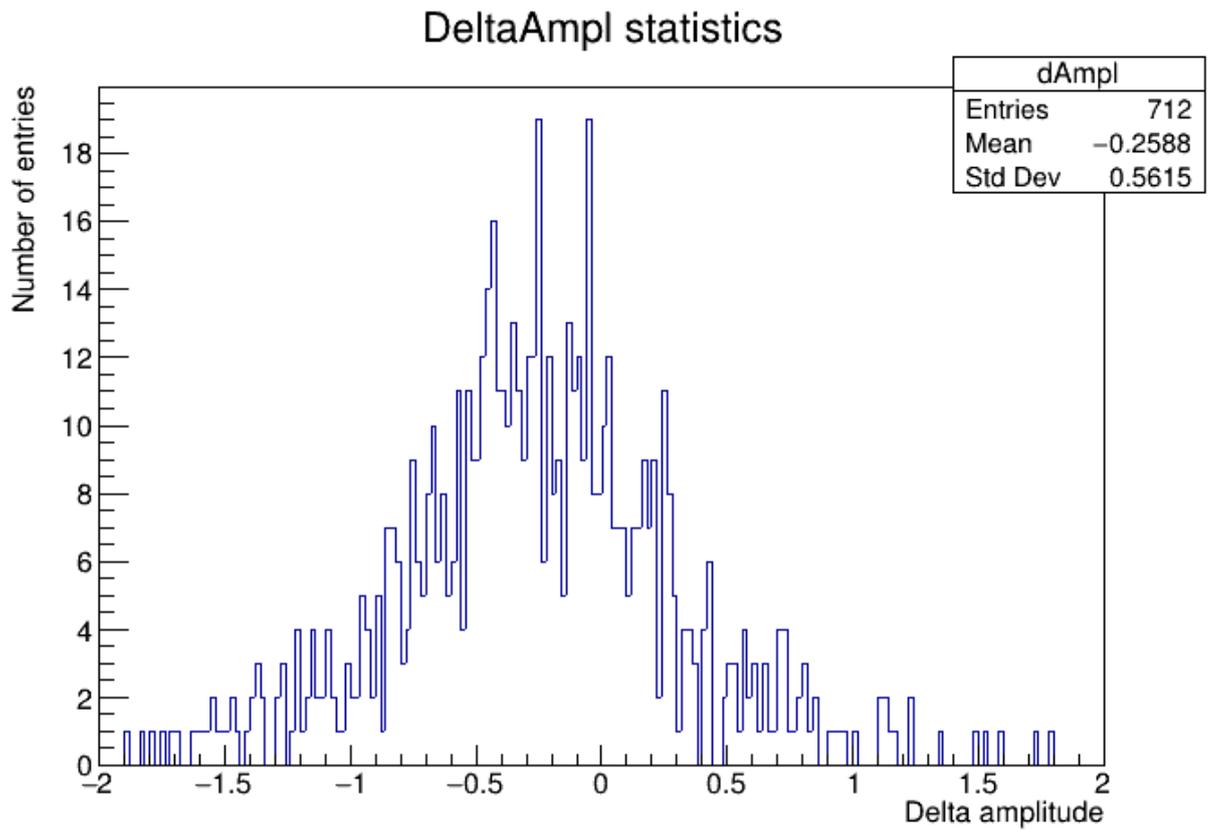


Рисунок 14 — Разность амплитуды двумя методами

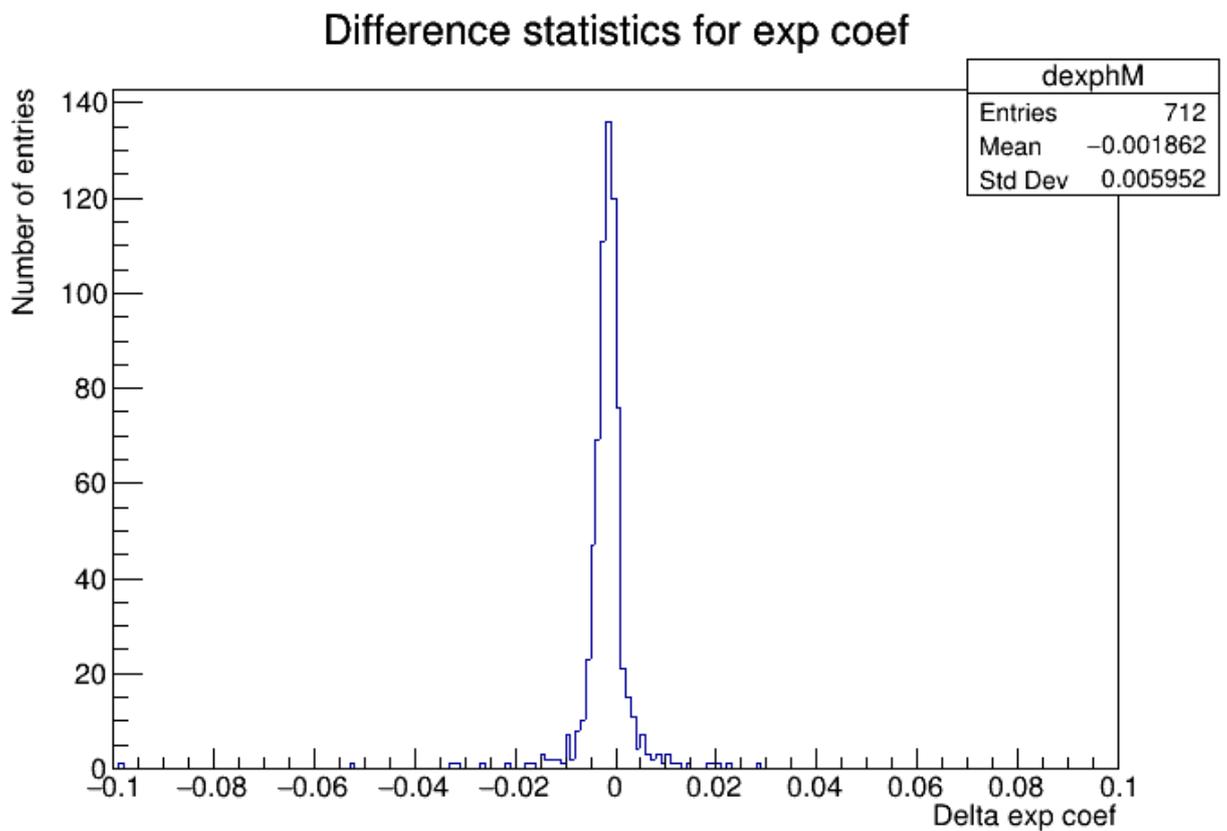


Рисунок 15 — Разность экспоненциального коэффициента двумя методами

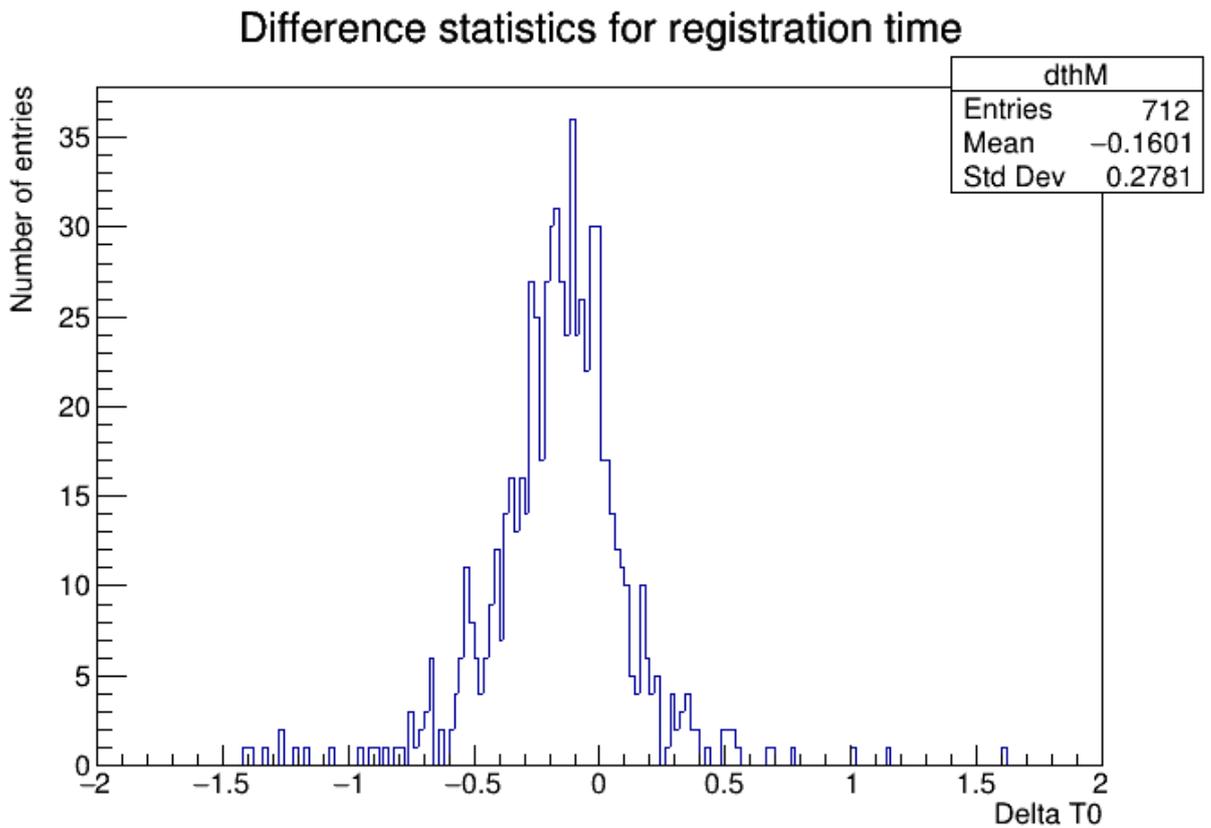


Рисунок 16 — Разность времени регистрации двумя методами

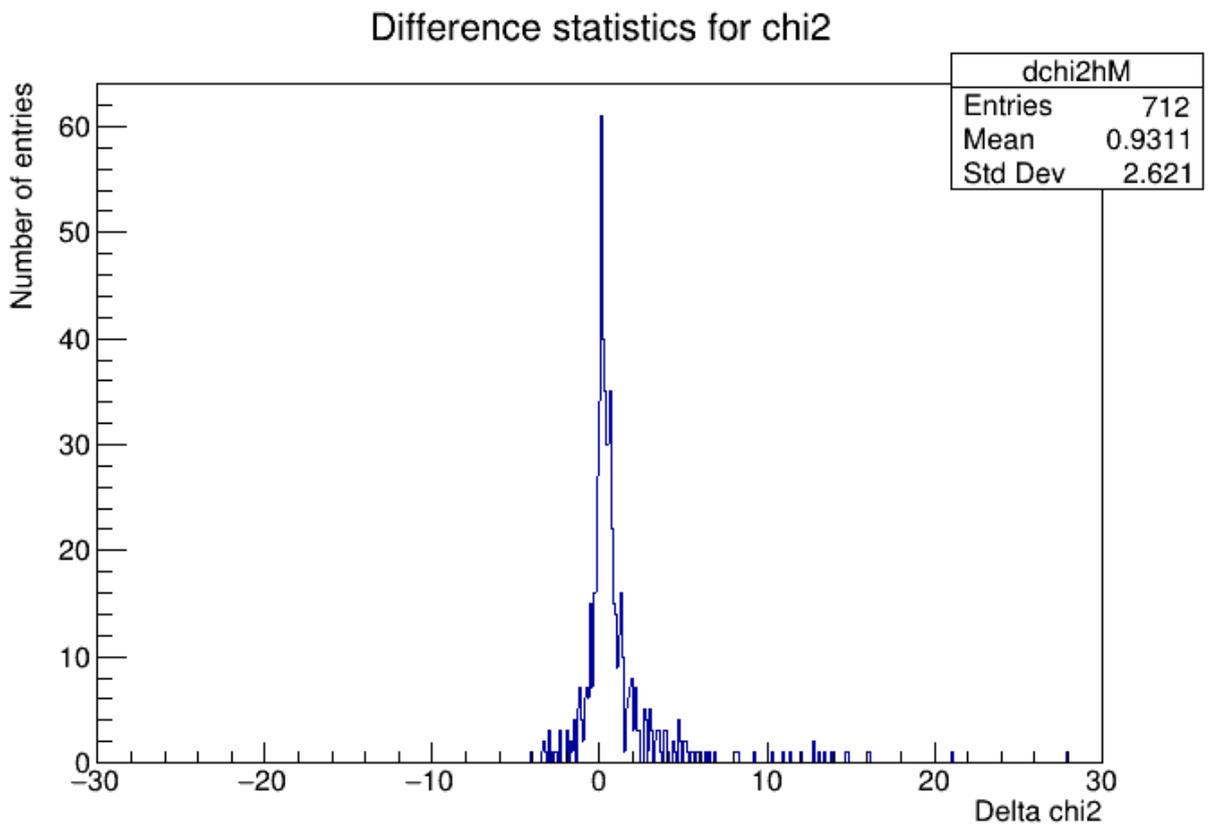


Рисунок 17 — Разность получаемого разными методами  $\chi^2$

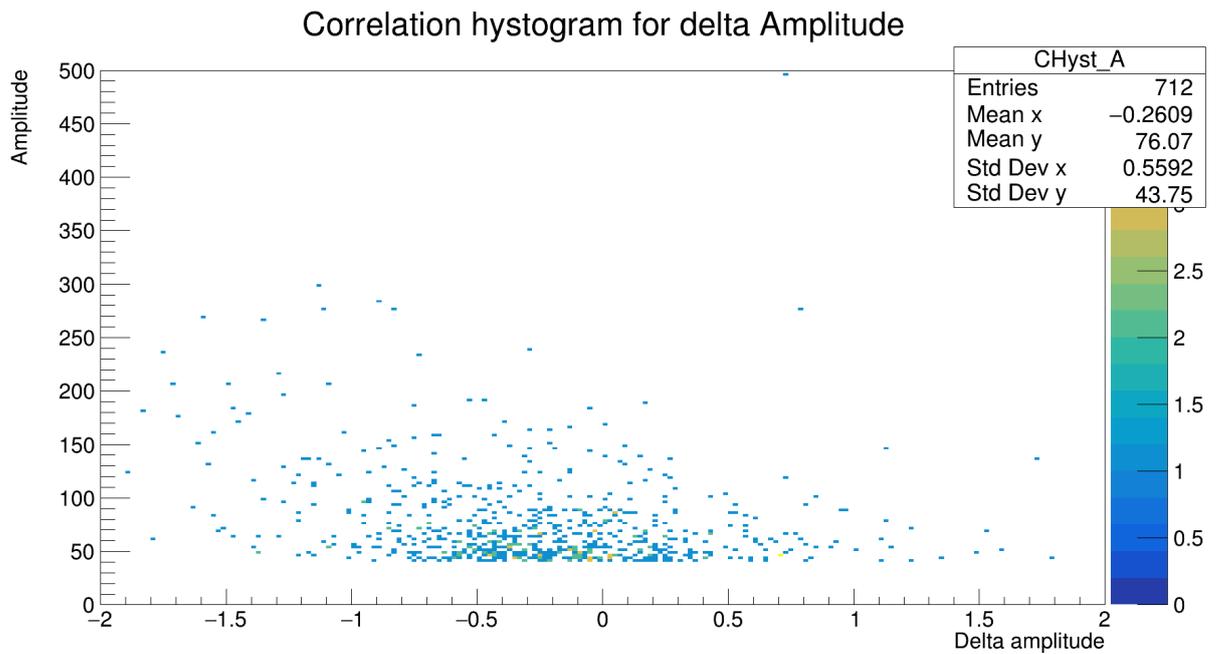


Рисунок 18 — Гистограмма распределения по разности амплитуды и амплитуде

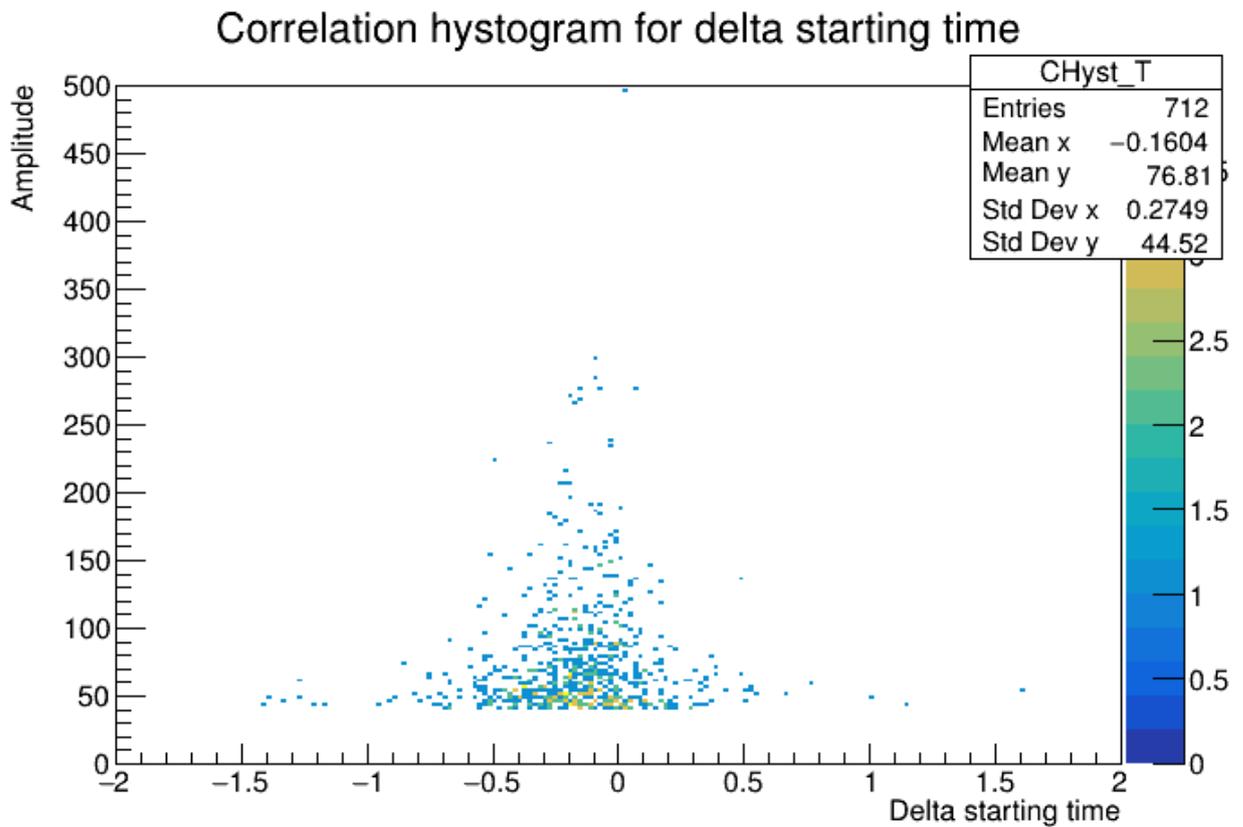


Рисунок 19 — Гистограмма распределения по разности времён регистрации и амплитуде

Sigma delta time from Ampl

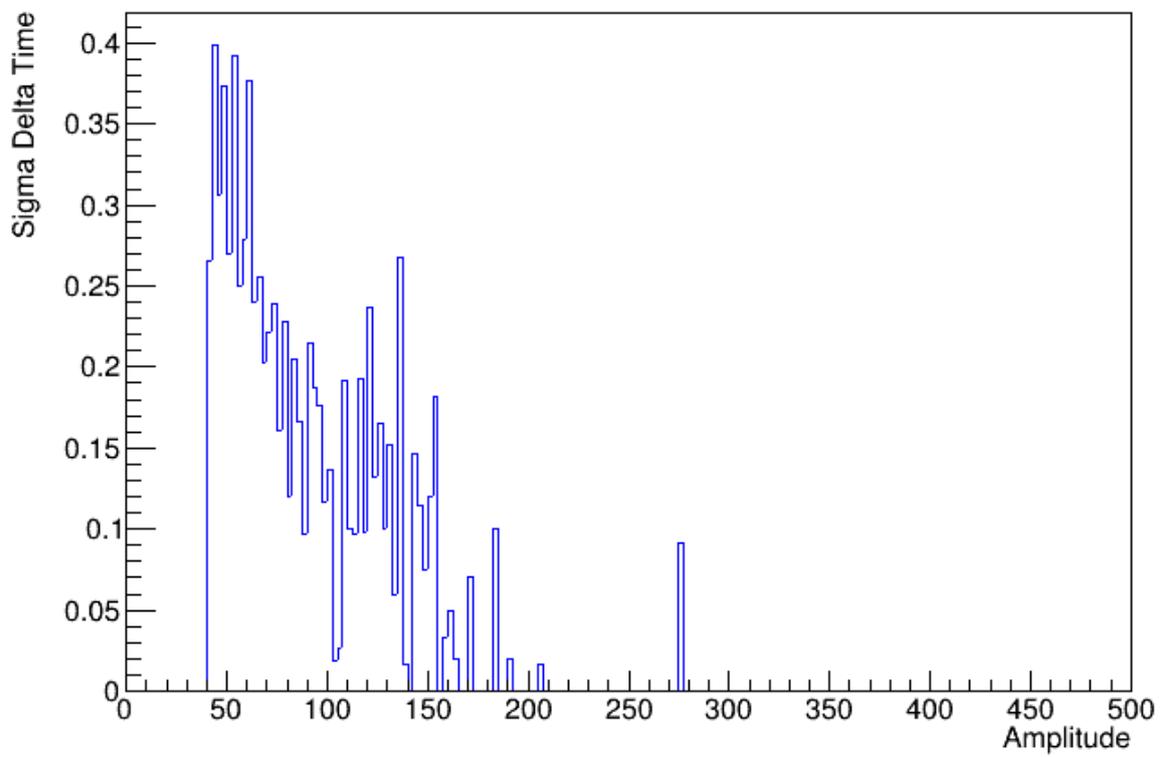


Рисунок 20 — Временное разрешение в зависимости от энергии

# Сравнение времени работы алгоритмов

Эта задача достаточно проста. Возможности языка C++ в виде пакета `time.h` позволяют вычислять время работы алгоритмов в тактах процессора. Поскольку применяться алгоритм будет вообще говоря вполне возможно на ЭВМ с иными характеристиками, то именно в них и надо вычислить сравнение.

Статистика по времени работы алгоритмов:

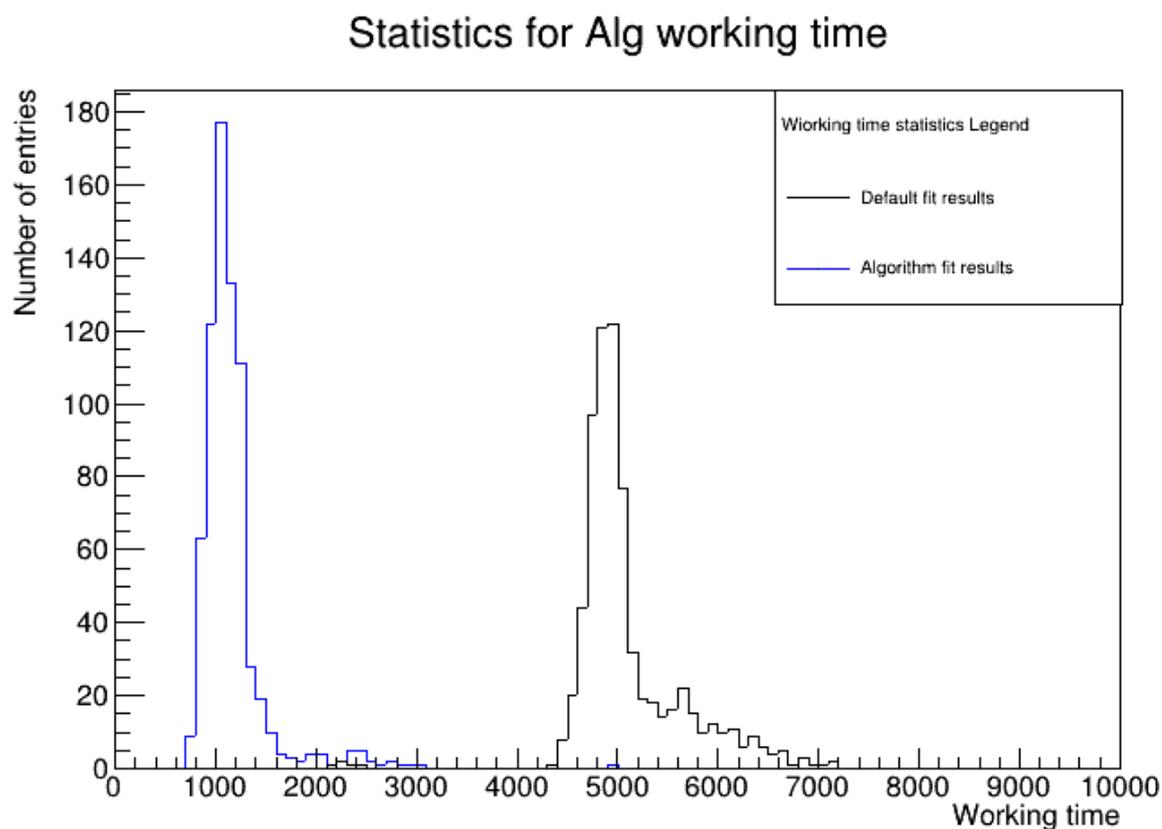


Рисунок 21 — Статистика времён работы алгоритмов

# Выводы

Представленные данные указывают, что несмотря на некоторые отклонения не выясненной природы, построенный алгоритм фитирует данные. Он имеет ошибку в вычислении, однако, она находится в допустимых пределах, то есть тех же, которые и были ранее. Полученная скорость работы недостаточна, что указывает на необходимость скорректировать исполнение, в частности, исполнить фитирование параболой иным образом, а также оптимизировать код, поскольку сведение задачи минимизации ФМП по 3 параметрам к минимизации по 1 параметру должна ускорять алгоритм существенно быстрее, чем это удаётся. Также следует отметить, что данная конкретная работа является частью более крупной работы, в которой подбирается оптимальное соотношение между точностью и скоростью вычисления параметров для калориметра PHOS в эксперименте ALICE.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. *Collaboration A.* [https://alice-collaboration.web.cern.ch/menu\\_projects/phos-cpv](https://alice-collaboration.web.cern.ch/menu_projects/phos-cpv). — 2020.
2. Analytical solutions in amplitude and time measurements from discrete sampling of pseudo-Gaussian signals / A. V. M.Yu. Bogolyubsky [et al.] // Nuclear Inst. and Methods in Physics Research, A. — 2021. — Vol. 1003, issue 1. — P. 165289.