

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

УДК 531.3, 539.1.05

ОТЧЕТ  
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ  
**ПРИМЕНЕНИЕ МЕТОДОВ ИСКУССТВЕННОГО  
ИНТЕЛЛЕКТА В ОБРАБОТКЕ ДАННЫХ ФИЗИЧЕСКОГО  
ЭКСПЕРИМЕНТА**

Научный руководитель

к. ф-м. н.

Студент

\_\_\_\_\_ А. С. Жемчугов

\_\_\_\_\_ К. Ю. Массалов

Москва 2023

# СОДЕРЖАНИЕ

|   |           |
|---|-----------|
| <b>Введение</b>   | <b>4</b>  |
| <b>1 Экспериментальная установка Spin Physics Detector на строящемся в ОИЯИ коллайдере NICA</b> | <b>5</b>  |
| 1.1 Конструкция детектора . . . . .   | 5         |
| 1.2 FARICH . . . . .  | 6         |
| <b>2 Алгоритмы распознавания колец</b>  | <b>7</b>  |
| 2.1 Разделение частиц по вторым моментам . . . . .  | 7         |
| 2.2 Итерационный метод . . . . .  | 7         |
| 2.3 Отношение функций максимального правдоподобия . . . . .                                     | 7         |
| 2.4 Метод масок . . . . .   | 7         |
| 2.5 Нечеткая кластеризация . . . . .  | 7         |
| 2.6 Преобразование Хафа . . . . .   | 8         |
| 2.7 COP . . . . .   | 8         |
| 2.8 Фитирование эллипса . . . . .   | 8         |
| <b>3 Моделирование детектора</b>  | <b>9</b>  |
| 3.1 GEANT4 . . . . .  | 9         |
| 3.2 Моделирование детектора и событий . . . . .   | 9         |
| <b>Заключение</b>   | <b>11</b> |
| <b>Список использованных источников</b>   | <b>12</b> |

## ВВЕДЕНИЕ

Относительный вклад видимой барионной материи во Вселенную составляет около 5% от общей массы. Хотя барионная материя представляется достаточно изученной, вопрос о том, почему нуклоны имеют именно такую структуру и свойства, остается открытым. Одним из главных нерешенных вопросов в квантовой хромодинамике (КХД) остается понимание структуры и основных свойств нуклонов исходя из динамики кварков и глюонов внутри них.

Нуклон обладает спином, который отвечает за его магнитный момент и другие свойства. Изучение спиновой структуры нуклонов является важной задачей. Центральной задачей является понимание того, как спин нуклона формируется из спинов и орбитальных моментов кварков и глюонов.

За последние 25 лет эксперименты по глубоко неупругому рассеянию (CERN, DESY, JLab, SLAC) и высокоэнергетические столкновения протонов (RHIC) были основными источниками информации о спиновых структурных функциях нуклонов. Наши знания о внутренней структуре нуклонов все еще ограничены, особенно в отношении глюонного вклада.

Spin Physics Detector — универсальная установка для исследования спиновой структуры нуклонов и других спиновых явлений. Установка будет использовать поляризованные пучки протонов и дейтронов. Проект основан на большом опыте работы с поляризованными пучками в ОИЯИ. Основная цель эксперимента - комплексное исследование неполяризованного и поляризованного глюонной составляющей нуклона. Измерения спина в SPD на коллайдере NICA откроют новые возможности для понимания спиновой структуры нуклона.

На установке планируется частота соударений до 4 МГц и число выходных каналов с детекторов  $\sim 700000$ . Таким образом, ожидается поток данных примерно 20 Гбайт/с или 200 Пбайт/год. Такой объем информации необходимо уменьшать, избавляясь от неинформативных событий и шумов. Для этого планируется использовать онлайн-фильтр, представляющий собой высокопроизводительную систему, которая будет включать в себя гетерогенные вычислительные платформы, аналогичные многим высокопроизводительным вычислительным кластерам. Для реконструкции и отбрасывания неинтересных событий можно использовать методы машинного обучения, поскольку проблема классических алгоритмов в том, что они имеют невысокую скорость и плохо параллелятся, в отличие от нейронных сетей. Проблема же нейронных сетей в сложности интерпретации их решений: никогда нельзя знать наверняка, почему сеть выдала именно такой результат, и насколько она права.

Таким образом, планируется использовать нейронные сети и классические алгоритмы одновременно: первые обеспечат скорость работы, вторые позволят проверять правильность результатов нейросетей.

Цель работы состоит в создании нейросетевых методов реконструкции сигналов с детектора FARICH, изучении их характеристик и сравнении их с классическими подходами. Для этого первоочередными задачами являются обзор существующих методов реконструкции сигналов с FARICH и создание программы моделирования черенковских колец для обучения нейросетей.

# 1 ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА SPIN PHYSICS ДЕТЕКТОР НА СТРОЯЩЕМСЯ В ОИЯИ КОЛЛАЙДЕРЕ NICA

Коллаборация Spin Physics Detector предлагает установить универсальный детектор во второй точке взаимодействия строящегося ускорителя NICA в Дубне. Этот детектор будет использоваться для изучения спиновой структуры протона и дейтрона и других спиновых эффектов с помощью поляризованных пучков протонов и дейтронов с энергией столкновения до 27 ГэВ и светимостью до  $10^{32}$   $\text{см}^{-2} \text{с}^{-1}$ . Эксперимент SPD в NICA позволит заполнить кинематический пробел между измерениями при низких энергиях на установках ANKE-COSY и SATURNE, измерениями при высоких энергиях на ускорителе тяжелых ионов и будущими экспериментами на Большом адронном коллайдере (БАК). Уникальность NICA заключается в возможности работы с пучками поляризованных дейтронов при этих энергиях.

## 1.1 КОНСТРУКЦИЯ ДЕТЕКТОРА

Экспериментальная установка SPD (Рис. 1) проектируется как универсальный детектор 4 $\pi$  с расширенными возможностями для отслеживания и идентификации частиц на основе современных технологий. Вершинный детектор (vertex detector) на основе кремния обеспечит разрешение по позиции вершин на уровне менее 100 микрон, необходимое для восстановления вторичных вершин от распадов D- мезонов. Система трекинга (straw tracker), помещенных в соленоидное магнитное поле с напряженностью до 1 Тесла на оси детектора, обеспечивает разрешение по поперечной импульсу  $\sigma_{p_T}/p_T \approx 2\%$  для частиц с импульсом 1 ГэВ/с. Время-пролетная система (time-of-flight system) с временным разрешением около 60 пикосекунд обеспечивает разделение  $3\sigma$  для пионов/каонов и для каонов/протонов до энергий 1.2 и 2.2 ГэВ соответственно. Использование черенковского детектора на базе аэрогеля — FARICH — поможет расширить этот диапазон вплоть до 6 ГэВ. Обнаружение фотонов обеспечивается электромагнитным калориметром (electromagnetic calorimeter) с энергетическим разрешением  $5\%/\sqrt{E}$ . Для идентификации мюонов планируется использовать мюонную систему (range system). Она также может использоваться как грубый адронный калориметр. За локальную поляриметрию и контроль светимости будет отвечать пара счетчиков пучков (beam-beam counter) и калориметров малого угла, или малоапертурных калориметров (zero-degree calorimeter).

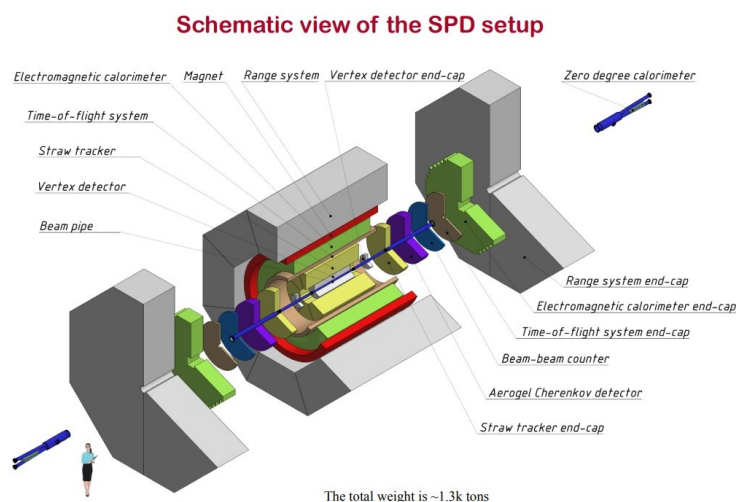


Рисунок 1 — Конструкция детектора

## 1.2 FARICH

FARICH (Focusing Aerogel Ring-Imaging CHerenkov detector) — черенковский детектор с аэрогелем, используемым в качестве рабочего тела. Черенковский детектор работает на основе эффекта Черенкова, который заключается в излучении света заряженными частицами, движущимися в среде со скоростью, превышающей фазовую скорость света в этой среде:  $\beta n > c$ . Это излучение имеет вид конуса, ось которого совпадает с направлением движения частицы, а угол раствора зависит от скорости частицы и показателя преломления среды. Излучение регистрируется массивом фотодетекторов, в результате чего на выходе получается срез конуса плоскостью. Если частица летела перпендикулярно детектору, получается набор точек, формирующих окружность, которую в дальнейшем на онлайн-фильтре надо будет реконструировать при помощи методов машинного обучения.

В качестве среды планируется использовать четыре слоя аэрогеля каждый со своим показателем преломления, что позволит улучшить разрешающую способность детектора [1].

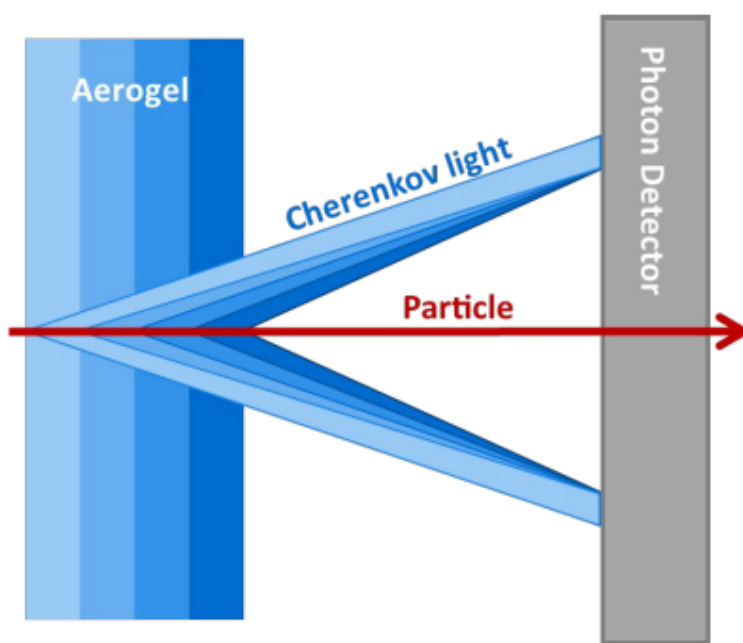


Рисунок 2 — Принципиальная схема детектора FARICH

## 2 АЛГОРИТМЫ РАСПОЗНАВАНИЯ КОЛЕЦ

### 2.1 РАЗДЕЛЕНИЕ ЧАСТИЦ ПО ВТОРЫМ МОМЕНТАМ

Простейшим методом идентификации является разделение по вторым моментам. Для каждого значения импульса второй центральный момент для кольца будет меняться в зависимости от типа частицы. Зная зависимость вторых моментов от импульса для разных частиц, можно определить их тип.

### 2.2 ИТЕРАЦИОННЫЙ МЕТОД

Для данного метода необходима дополнительная информация о предположительных центрах колец, которую можно получить независимо, например, в результате экстраполяции трека из трекового детектора.

Принцип работы следующий: при инициализации рассчитывается расстояние от каждого хита до каждого трека. Полученными значениями с весом 1 заполняют гистограммы (по числу треков). Затем, на каждой итерации обновляют веса хитов по формуле:

$$w_k = \frac{y_k}{\sum_i y_i}, \quad (1)$$

где  $w_k$  — вес хита в  $k$ -той гистограмме,  $y_k$  — содержание ячейки гистограммы, в которую попал данный хит. В результате получаются пики, которые соответствуют центрам колец [2].

### 2.3 ОТНОШЕНИЕ ФУНКЦИЙ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Данный метод также требует дополнительной информации о центрах колец. На основании информации о конструкции детектора и параметрах трека делается предсказание положения черенковского фотона при гипотезе о типе частицы. На основе предсказания и наблюдения строится функция правдоподобия, а затем выбирается функция с максимальным правдоподобием [2].

### 2.4 МЕТОД МАСОК

На предварительно центрированное изображение накладывается трафарет с известными параметрами. Наложение заключается в сложении пикселей маски и изображения по модулю два. Чем больше нулей в результате, тем больше данная маска соответствует изображению [2].

### 2.5 НЕЧЕТКАЯ КЛАСТЕРИЗАЦИЯ

Метод заключается в минимизации следующего функционала:

$$J = \sum_{i=1}^C \sum_{j=1}^N u_{ij} d_{ij}^2, \quad (2)$$

где  $u_{ij}$  — весовой коэффициент, показывающий вероятность принадлежности хита  $i$  к кластеру  $j$  ( $\sum_j u_{ij} = 1$ ),  $d_{ij}$  — расстояние между  $i$  хитом и  $j$  кластером,  $C$  — количество колец,  $N$  — количество хитов.

В данном методе необходимо также знать центры колец и их количество [2].

## 2.6 ПРЕОБРАЗОВАНИЕ ХАФА

Изначально преобразование Хафа предназначено для распознавания прямых и заключается в идее перейти из пространства координат в пространство параметров и нахождения там точки с наибольшей «интенсивностью», т.е. точки с наибольшим числом пересечений прямых. Полученная точка в пространстве параметров и есть параметры, описывающие прямую, проходящую через исходные точки.

Затем данное преобразование можно обобщить и на случай окружностей, перейдя в трехмерное пространство параметров [3].

## 2.7 СОР

Алгоритм СОР (Chernov-Ososkov-Pratt) позволяет по точкам оценить параметры окружности [4].

Для подгонки окружности методом наименьших квадратов необходимо минимизировать функционал:

$$L(a, b, R) = \sum_{i=1}^N (\sqrt{(x_i - a)^2 + (y_i - b)^2} - R)^2, \quad (3)$$

где  $a, b$  —  $x$  и  $y$  координаты центра окружности,  $R$  — ее радиус,  $x_i$  и  $y_i$  — координаты  $i$ -й точки,  $N$  — число точек.

Однако, такой метод требует достаточно много времени для вычисления. Для ускорения работы алгоритма можно использовать следующий функционал:

$$M(a, b, c) = \sum_{i=1}^N ((x_i - a)^2 + (y_i - b)^2 - R^2)^2 \quad (4)$$

Минус такого подхода в том, что при измерениях, не покрывающих всю окружность, результат может оказаться довольно далеко от истины.

В алгоритме СОР предлагается использовать функционал:

$$M(a, b, c) = \sum_{i=1}^N [((x_i - a)^2 + (y_i - b)^2 - R^2)^2 / R^2] \quad (5)$$

Нахождение минимума такого функционала эквивалентно минимизации (3), однако вычислительно менее затратно.

## 2.8 ФИТИРОВАНИЕ ЭЛЛИПСА

Для фитирования эллипса предлагается использовать функционал вида [5]:

$$L(x_{F_1}, y_{F_1}, x_{F_2}, y_{F_2}, a) = \sum_{i=1}^N (\sqrt{(x_i - x_{F_1})^2 + (y_i - y_{F_1})^2} + \sqrt{(x_i - x_{F_2})^2 + (y_i - y_{F_2})^2} - 2a)^2, \quad (6)$$

где  $x_{F_1}, y_{F_1}, x_{F_2}, y_{F_2}$  — координаты  $x$  и  $y$  первого и второго фокуса эллипса соответственно,  $a$  — длина большой полуоси.

Минимизация осуществляется при помощи градиентного спуска.

## 3 МОДЕЛИРОВАНИЕ ДЕТЕКТОРА

### 3.1 GEANT4

GEANT4 — это программный инструмент на C++ для моделирования экспериментов в области физики элементарных частиц. Он разрабатывается и поддерживается Европейской организацией по ядерным исследованиям (CERN) [6–8].

GEANT4 используется для моделирования прохождения частиц через материю, включая взаимодействия, которые происходят при столкновении частиц с атомами в материале. Эта информация может быть использована для проектирования экспериментов, понимания результатов этих экспериментов и создания симуляций будущих экспериментов.

В данной работе пакет GEANT4 использовался для моделирования детектора FARICH.

### 3.2 МОДЕЛИРОВАНИЕ ДЕТЕКТОРА И СОБЫТИЙ

Модель детектора (Рис. 3) — два диска радиусом 60 см. Первый диск состоит из четырех слоев аэрогеля на основе диоксида кремния (94% по массе) с небольшой примесью диоксида циркония (6% по массе) с изменяющимся показателем преломления: 1.022, 1.024, 1.027, 1.03. Толщины слоев 7 мм, 6.7, 6.3 мм, 6 мм соответственно. Второй — чувствительный детектор, который регистрирует черенковские фотоны в диапазоне примерно от 400 нм до 450 нм — в области чувствительности кремниевых фотодетекторов. Расстояние между аэрогелем и фотоприемником — 20 см.

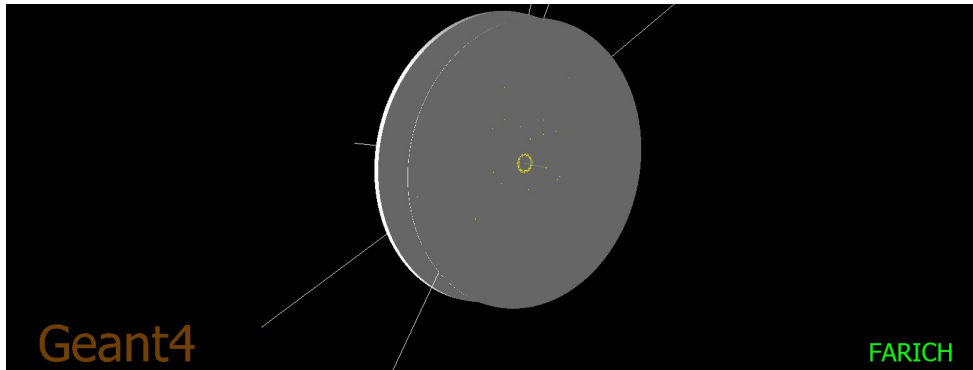


Рисунок 3 — Моделирование детектора FARICH

Во время моделирования для каждого из четырех типов частиц при пяти значениях импульсов (1-5 ГэВ) запускалось 100 событий. В каждом событии частица падала перпендикулярно детектору в точке (0,0). Координаты и энергия черенковских фотонов из диапазона 400-450 нм сохранялись. Затем строились гистограммы для числа фотонов в каждом событии (Рис. 4), а также вторые центральные моменты для колец по формуле:

$$\mu_2 = \frac{1}{n} \sum_{i=1}^n (\sqrt{x_i^2 + y_i^2}) - \frac{1}{n} \sum_{j=1}^n (\sqrt{x_j^2 + y_j^2}) \quad (7)$$

где  $x_i$  и  $y_i$  — координаты  $i$ -той точки,  $n$  — число точек в кольце.

В результате моделирования были построены спектры фотонов, гистограммы для радиусов колец и вторых моментов для  $e^-$ ,  $\pi^+$ ,  $K^+$ ,  $p^+$  для импульсов от 1 до 5 ГэВ.

Из приведенных графиков видно, что по моментам возможно разделить электроны и пионы до импульса примерно в 1-2 ГэВ. Также неплохо разделяются каоны и протоны до импульсов в 5 ГэВ. Однако, требуется дальнейшее совершенствование метода для пионов и каонов.



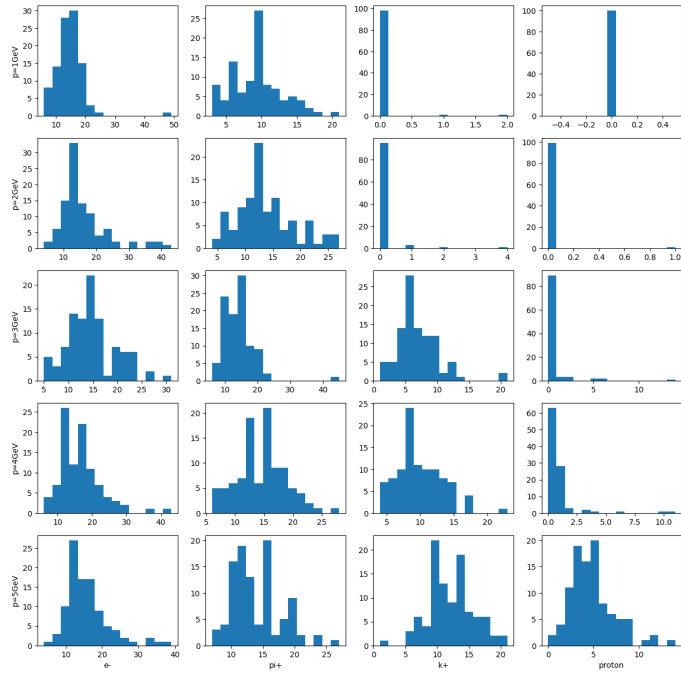


Рисунок 4 — Энергетические спектры для разных импульсов и частиц

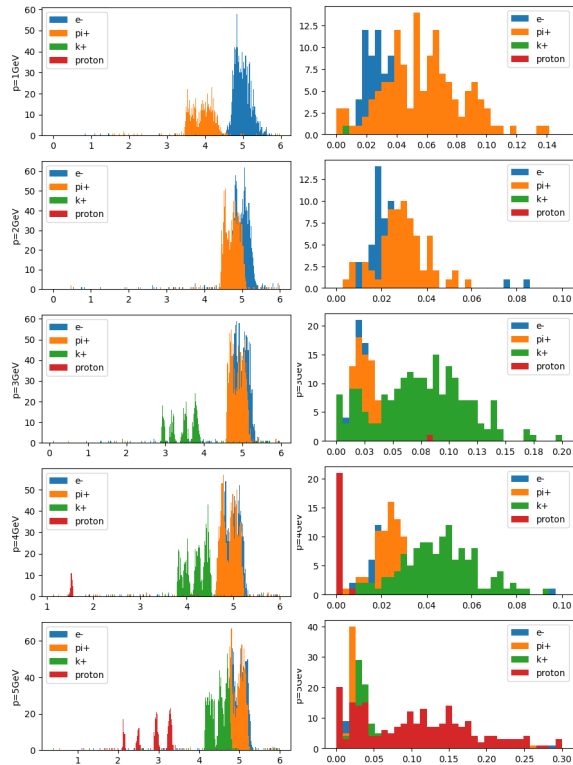


Рисунок 5 — Радиусы(слева) и вторые моменты(справа) для разных импульсов и частиц. Радиусы взяты из истинной информации Монте-Карло генерирования.

## ЗАКЛЮЧЕНИЕ

В ходе работы сделан обзор основных методов реконструкции колец в RICH детекторах.

При помощи программного пакета GEANT4 смоделированы события пролета разных частиц при разных импульсах через детектор FARICH. Построены гистограммы для радиусов и вторых центральных моментов получившихся колец. Из получившихся гистограмм сделан вывод о возможности разделения  $\pi^+$ ,  $K^+$ ,  $p^+$ , но эффективность такого метода невысока.

В дальнейшей работе планируется уточнение модели детектора с более точным учетом эффективности детектирования черенковских фотонов, изучение эффективности классических алгоритмов восстановления с учетом различных углов падения частиц на радиатор FARICH, а также разработка и обучение нейросети для идентификации частиц.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. RICH with multiple aerogel layers of different refractive index / S. Korpar [и др.] // Nucl. Instrum. Meth. A / под ред. J. Engelfried, G. Paic. — 2005. — Т. 553.
2. *Лебедев С. А.* Математическое обеспечение для реконструкции колец черенковского излучения и идентификации электронов в RICH детекторе эксперимента CBM : дис. . . . канд. / Лебедев С. А. — Объединённый институт ядерных исследований, 2011.
3. *Hough V P. C.* Method and means for recognizing complex patterns. — 1962.
4. *Chernov N. I., Ososkov G. A.* Effective Algorithms of Circle Fitting // Comput. Phys. Commun. — 1984. — Т. 33. — С. 329—333.
5. *Lebedev S. A., Ososkov G. A.* Fast algorithms for ring recognition and electron identification in the CBM RICH detector // Phys. Part. Nucl. Lett. — 2009. — Т. 6. — С. 161—176.
6. Recent developments in Geant4 / J. Allison [и др.] // Nucl. Instrum. Meth. A. — 2016. — Т. 835. — С. 186—225.
7. Geant4 developments and applications / J. Allison [и др.] // IEEE Transactions on Nuclear Science. — 2006. — Т. 53, № 1. — С. 270—278.
8. GEANT4—a simulation toolkit / S. Agostinelli [и др.] // Nucl. Instrum. Meth. A. — 2003. — Т. 506. — С. 250—303.