

Выпускная квалификационная работа
Монте-Карло моделирование формы
импульса и дискриминация событий при
помощи нейронных сетей в эксперименте
GERDA

Зарецкий Н.Д.
Научный руководитель: Гробов А.В.

Национальный исследовательский ядерный университет «МИФИ»

29 июня 2021 г.

Введение

Цель работы:

- анализ данных, полученных на полукоаксиальных детекторах эксперимента GERDA, с использованием нейросети и дальнейшее применение нейросети для дискриминации фоновых и сигнальных событий.
- Монте-Карло моделирование импульсов с помощью генеративно-состязательных сетей (GAN).

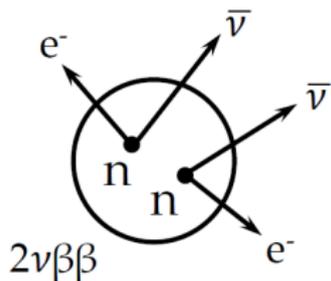


Рисунок 1 – Схема $2\nu\beta\beta$

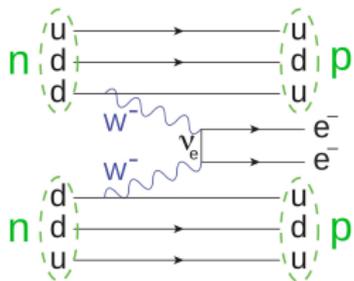


Рисунок 2 – Диаграмма $0\nu\beta\beta$

Эксперимент GERDA

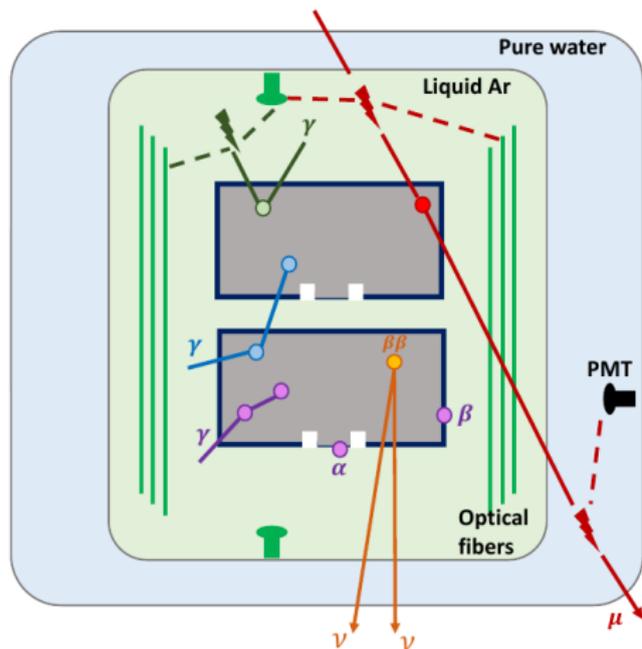


Рисунок 3 – Схема эксперимента GERDA

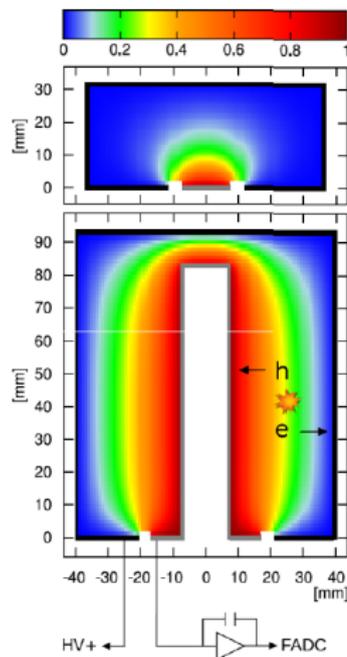


Рисунок 4 – Схема полукоаксиальных и BEGe детекторов

Данные с эксперимента GERDA

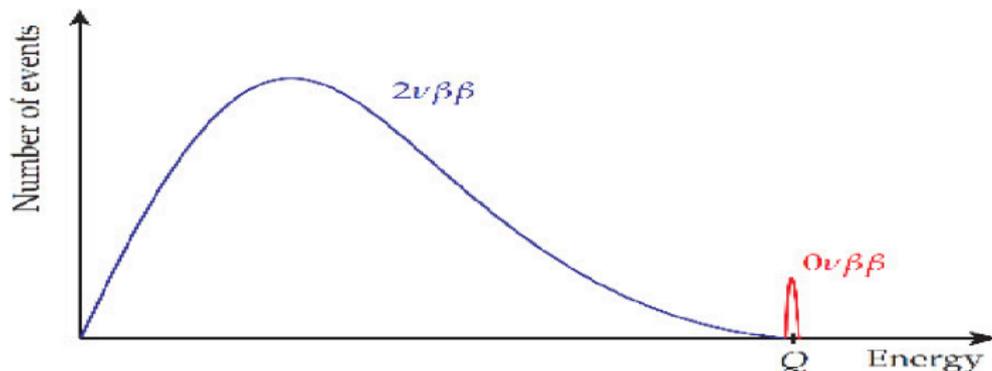


Рисунок 5 – Спектр энергии двух электронов

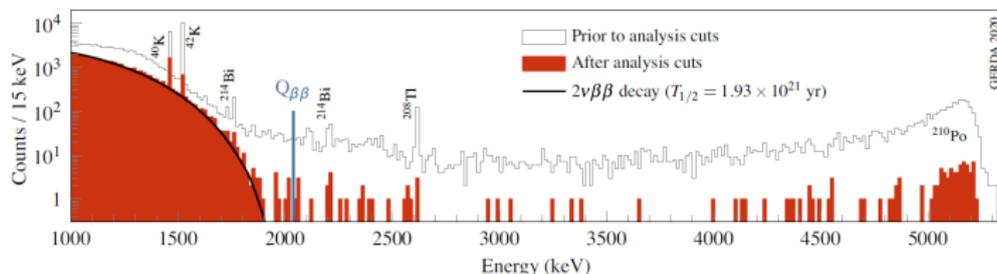


Рисунок 6 – Спектр энергии с эксперимента GERDA

Формы импульсов для полукоаксиальных детекторов

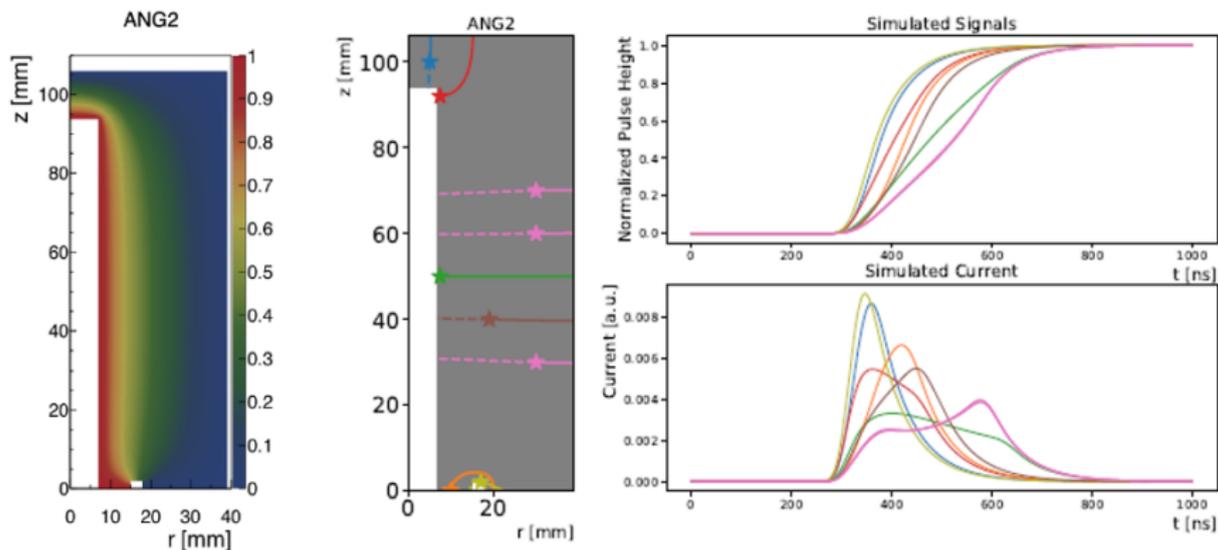


Рисунок 7 – Возможные формы импульсов для разных мест выделения энергии в полукоаксиальном детекторе

Калибровки

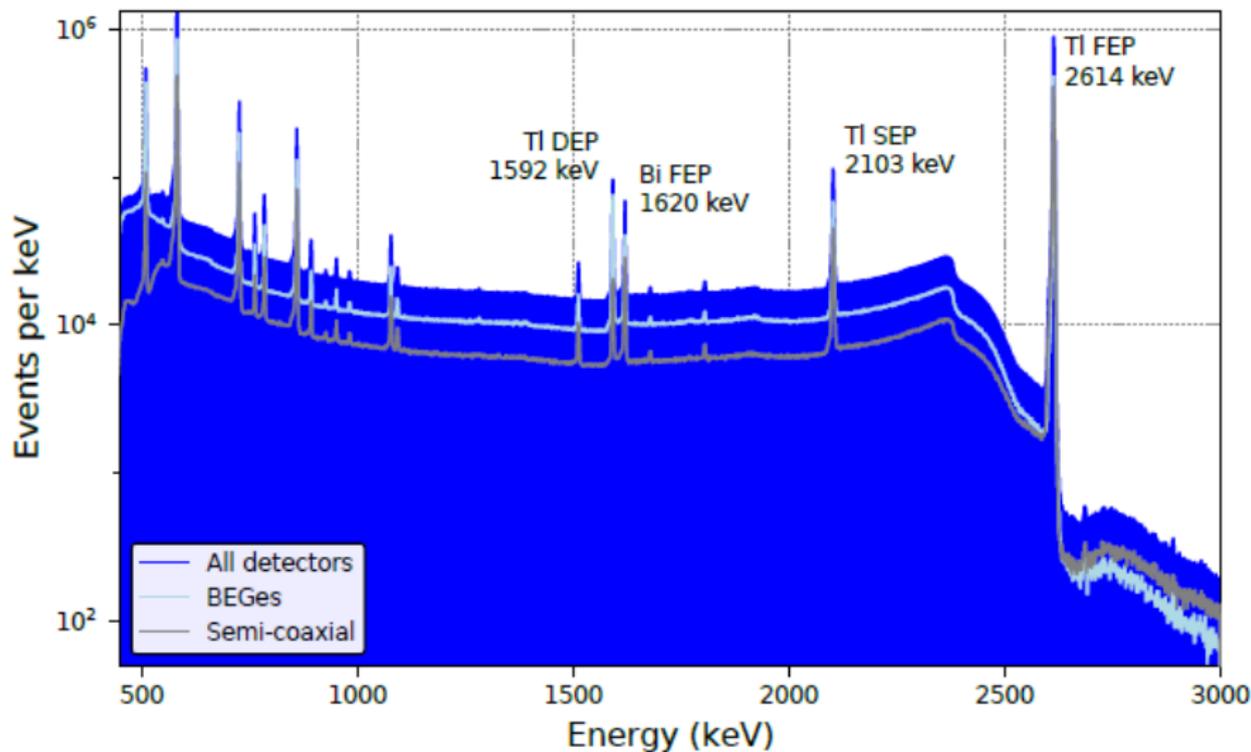


Рисунок 8 – Энергетический спектр ^{228}Th

Анализ калибровочных данных

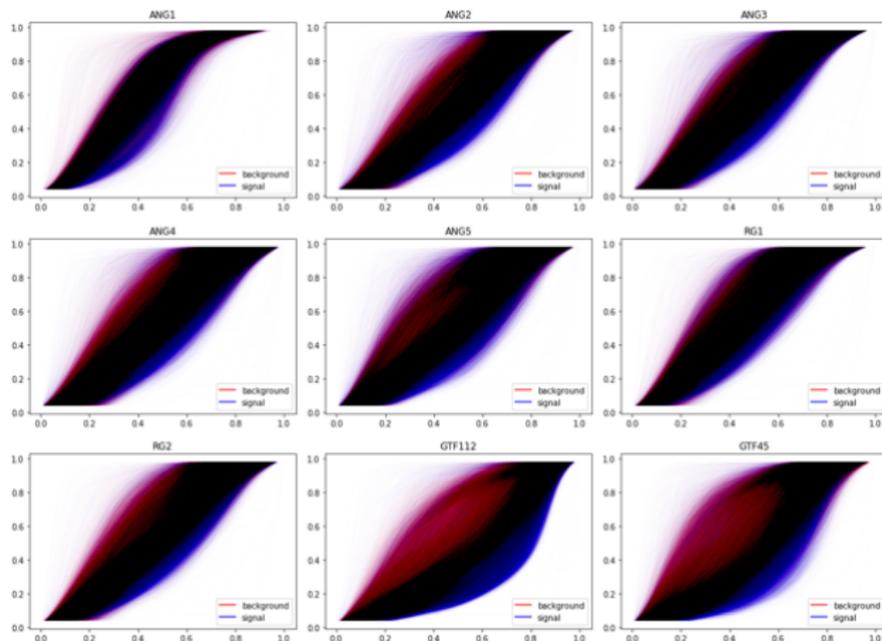


Рисунок 9 – Формы импульсов, полученные в полукоаксиальных детекторах эксперимента GERDA. По горизонтальной оси - время, нормированное на длительность импульса

Анализ калибровочных данных

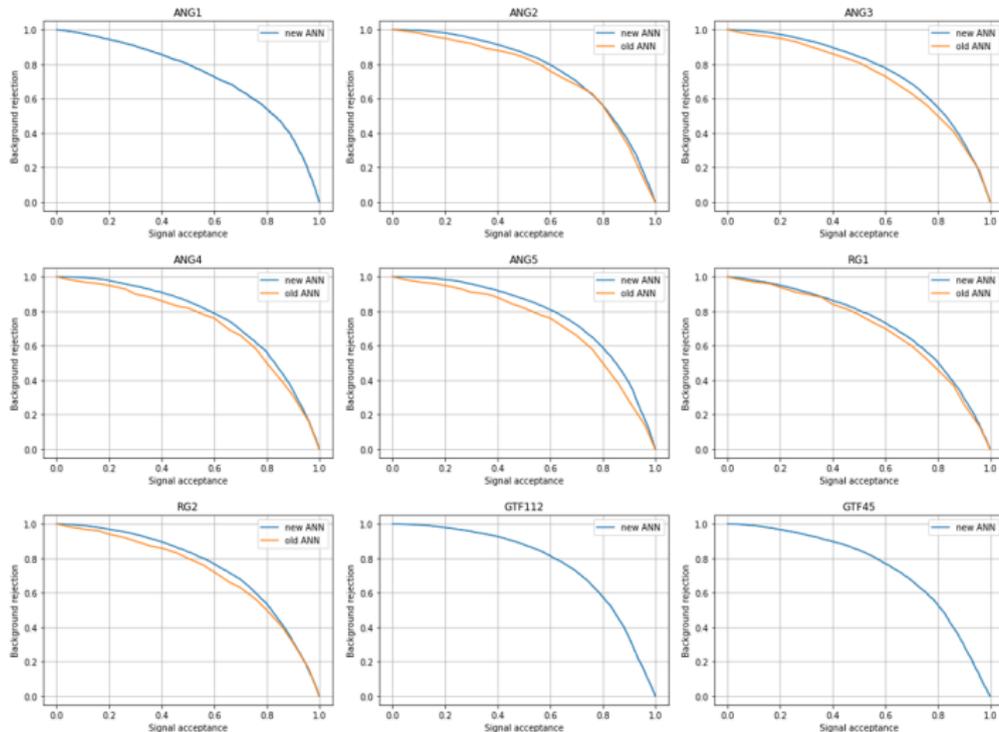


Рисунок 10 – ROC-кривые для каждого детектора при обучении и проверке нейросети на калибровочных данных

Анализ калибровочных данных

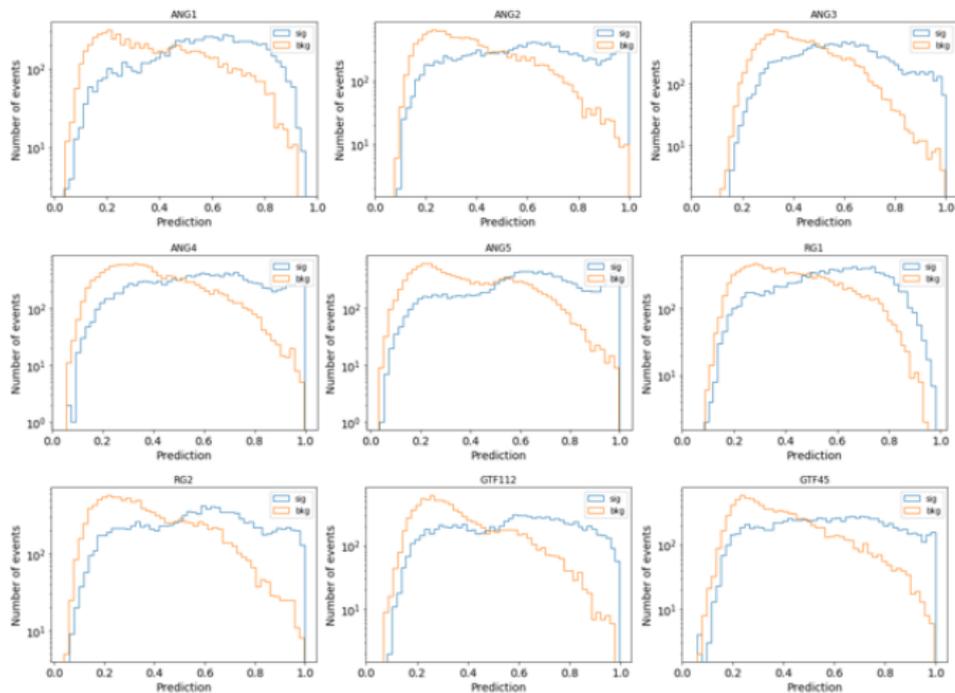


Рисунок 11 – Разделение сигнальных и фоновых событий для каждого детектора при обучении и проверке нейросети на калибровочных данных

Анализ Монте-Карло данных

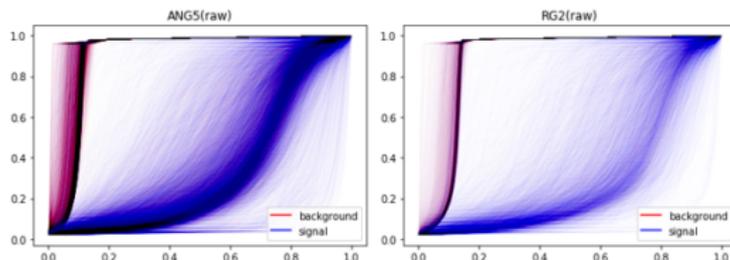


Рисунок 12 – Формы импульсов для Монте-Карло данных. По горизонтальной оси - время, нормированное на длительность импульса

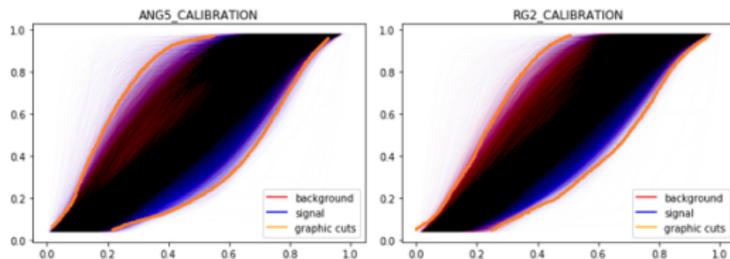


Рисунок 13 – Формы импульсов, полученные на эксперименте GERDA, с границей области графического критерия отбора. По горизонтальной оси - время, нормированное на длительность импульса

Анализ Монте-Карло данных

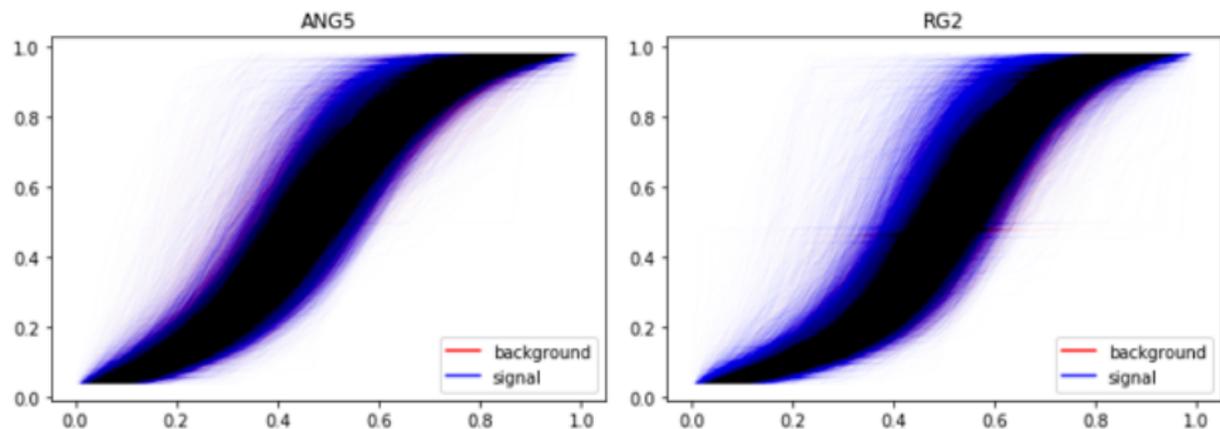


Рисунок 14 – Формы импульсов для Монте-Карло данных после преобразований. По горизонтальной оси - время, нормированное на длительность импульса

Анализ Монте-Карло данных

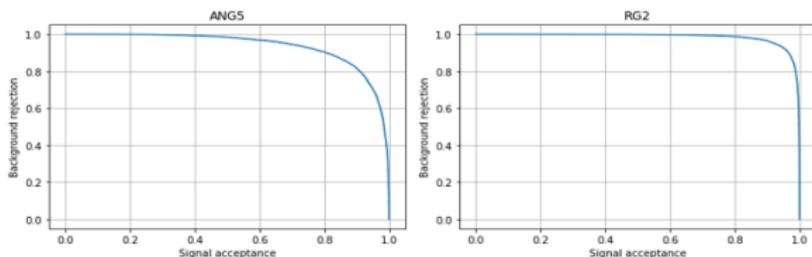


Рисунок 15 – ROC-кривые при обучении и проверке нейросети на Монте-Карло данных

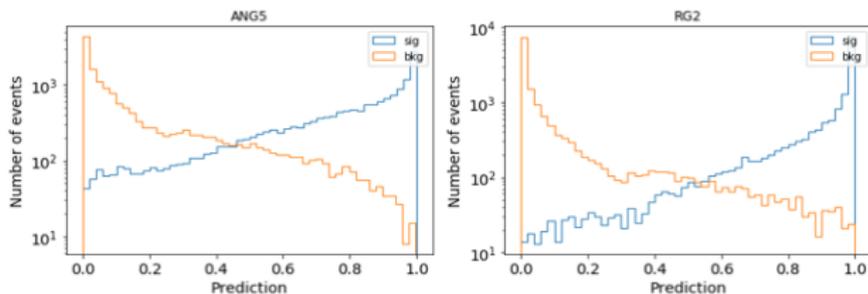


Рисунок 16 – Разделение сигнальных и фоновых событий при обучении и проверке нейросети на Монте-Карло данных

Анализ Монте-Карло данных

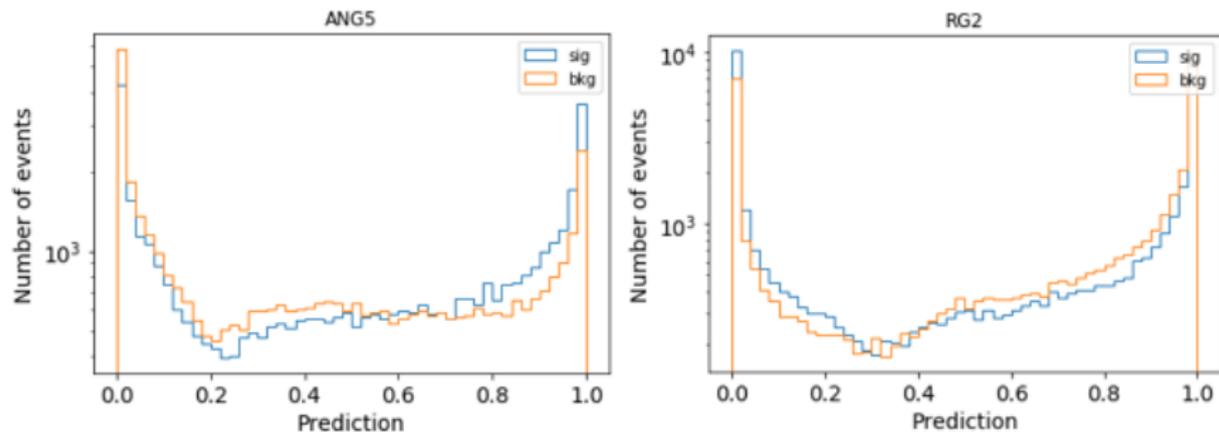


Рисунок 17 – Разделение сигнальных и фоновых событий при обучении нейросети на Монте-Карло данных и ее проверке на калибровочных данных

Анализ Монте-Карло данных

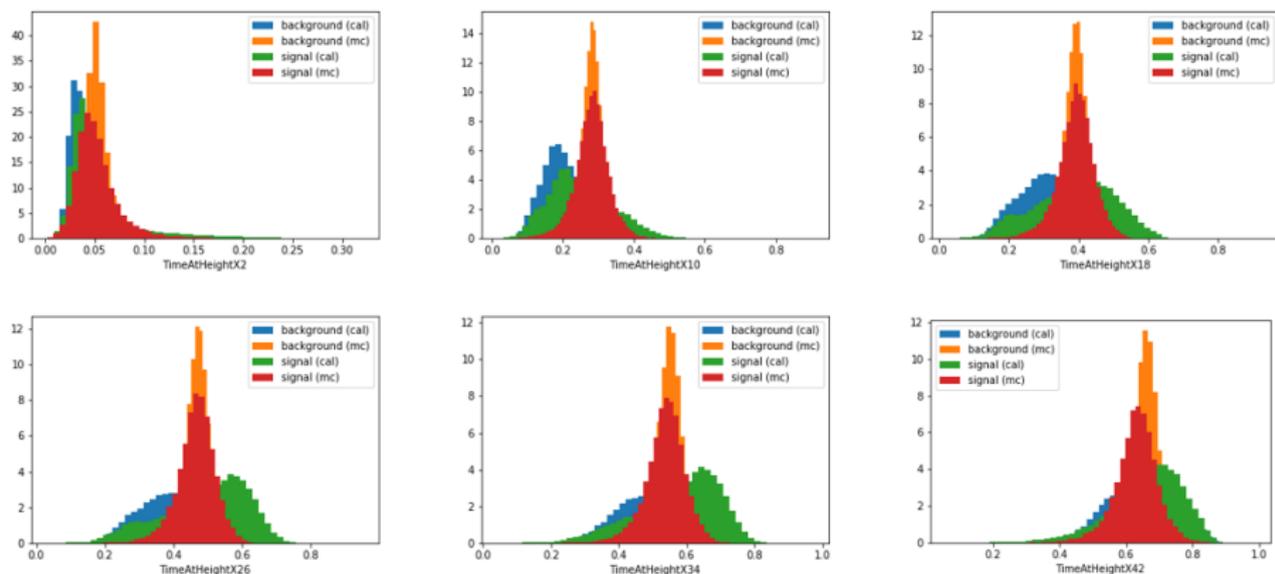


Рисунок 18 – Гистограммы некоторых параметров TimeAtHeightX в Монте-Карло данных для детектора ANG5

Монте-Карло моделирование формы импульса с помощью сетей GAN

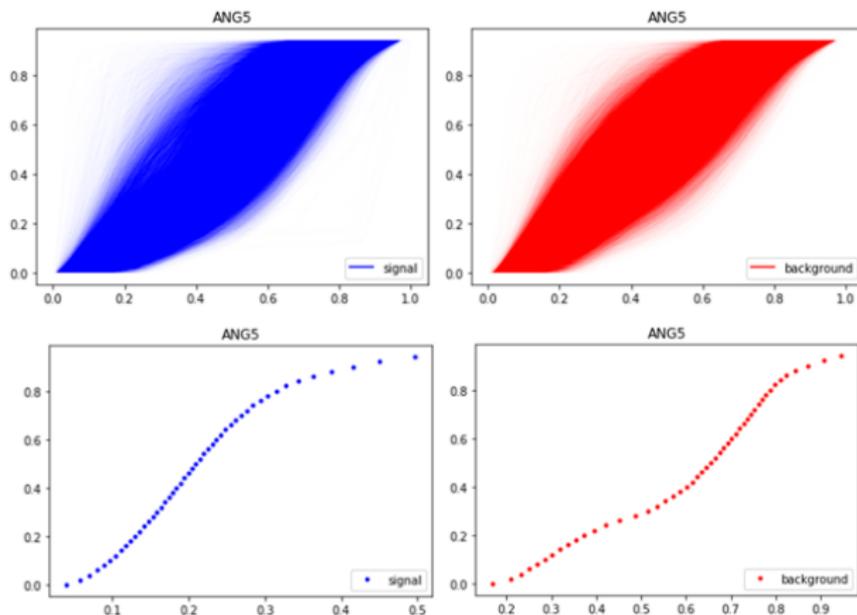


Рисунок 19 – Пример форм сигнальных и фоновых импульсов, полученных на эксперименте GERDA. По горизонтальной оси - время, нормированное на длительность импульса

Монте-Карло моделирование формы импульса с помощью сетей GAN

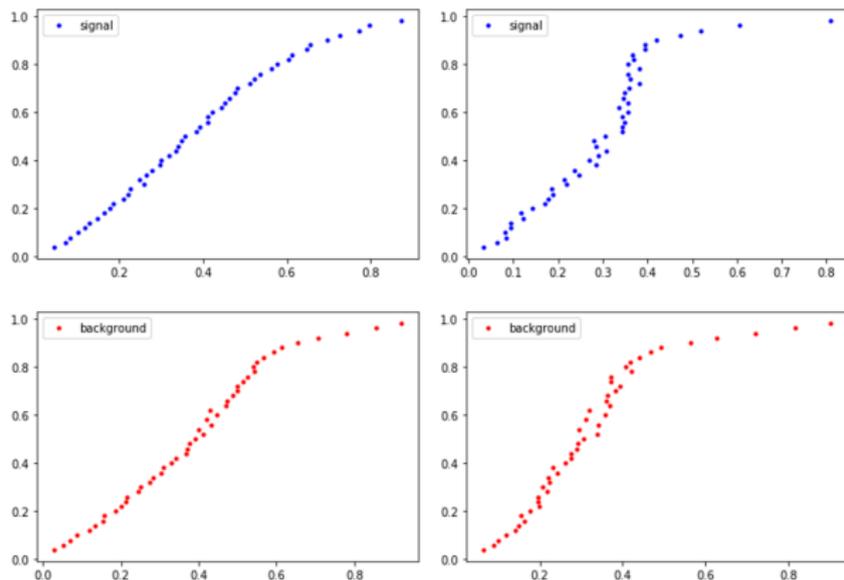


Рисунок 20 – Пример форм сигнальных и фоновых импульсов, сгенерированных перцептроном, обученным на стандартизованных данных

Монте-Карло моделирование формы импульса с помощью сетей GAN

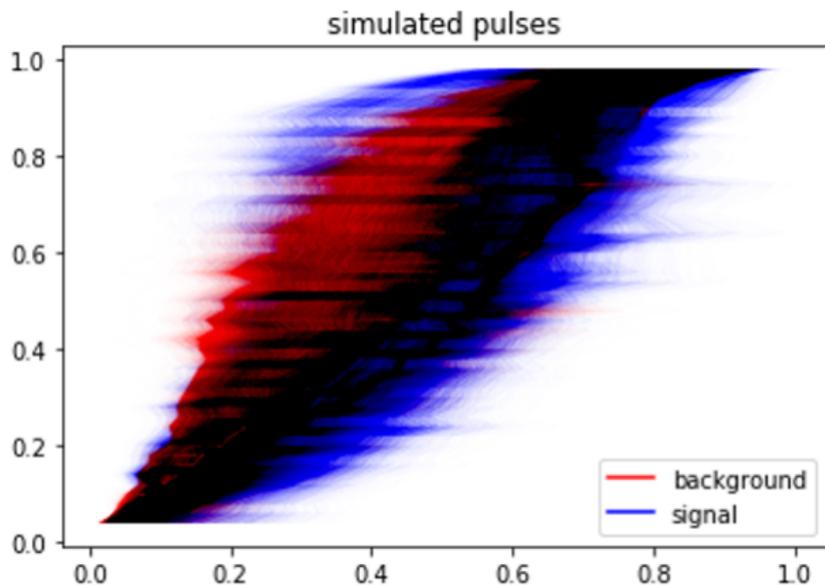


Рисунок 21 – Импульсы, сгенерированные перцептроном, обученным на стандартизованных данных

Монте-Карло моделирование формы импульса с помощью сетей GAN

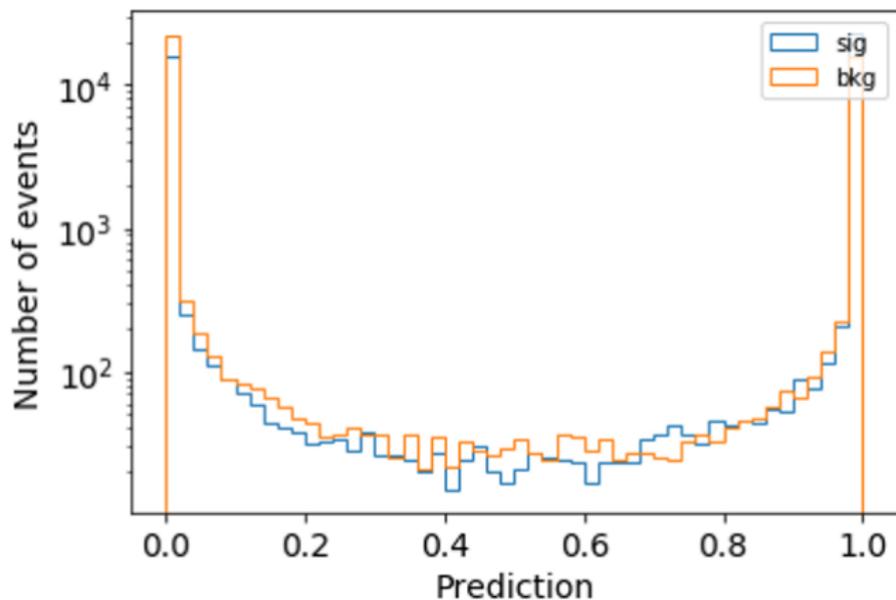


Рисунок 22 – Разделение сигнала и фона для детектора ANG5 при обучении нейросети на сгенерированных данных и ее проверке на калибровочных данных

Монте-Карло моделирование формы импульса с помощью сетей GAN

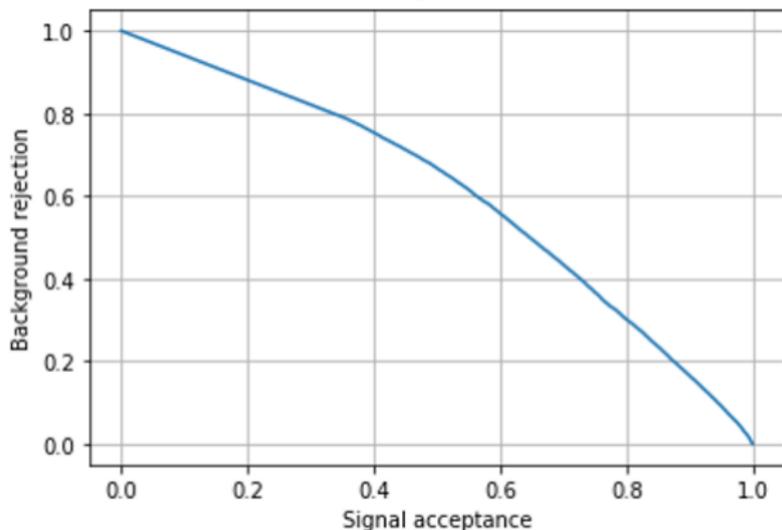


Рисунок 23 – ROC-кривая для детектора ANG5 при обучении нейросети на сгенерированных данных и ее проверке на калибровочных данных

Заключение

В ходе работы были сделаны следующие выводы:

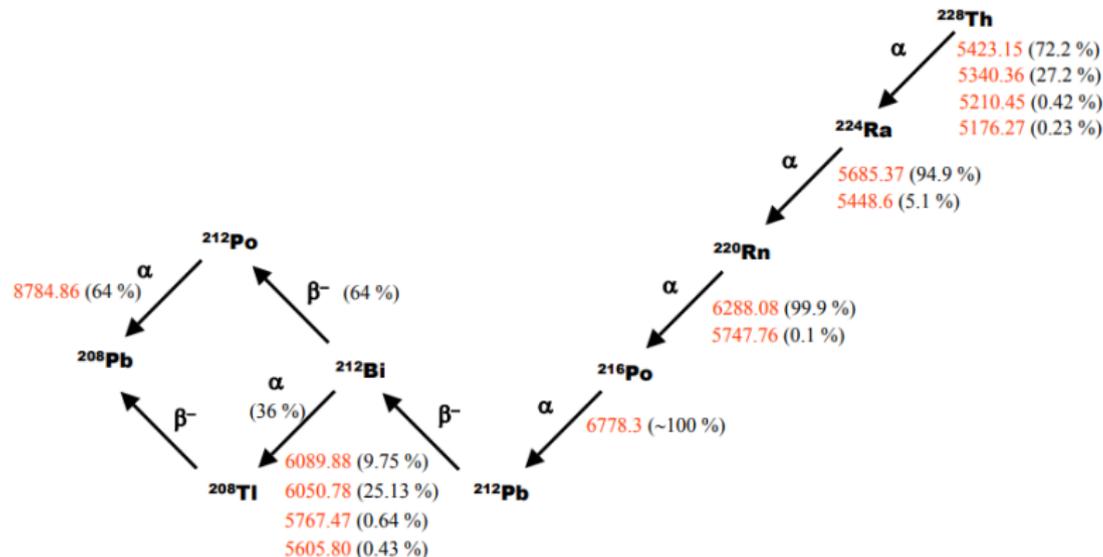
- результаты по разделению сигнальных и фоновых событий в **калибровочных** данных детектора GERDA были улучшены по сравнению с предшествующим анализом, однако разделение по-прежнему остается трудновыполнимым.
- анализ имеющихся **Монте-Карло** данных свидетельствует об ошибках в Монте-Карло моделировании.
- нейросеть для бинарной классификации, обученная на синтезированных данных, работает при ее проверке на калибровочных данных, поэтому при условии улучшения работы сети-генератора можно использовать импульсы, синтезированные сетью GAN, в качестве Монте-Карло данных.

Дополнительные слайды

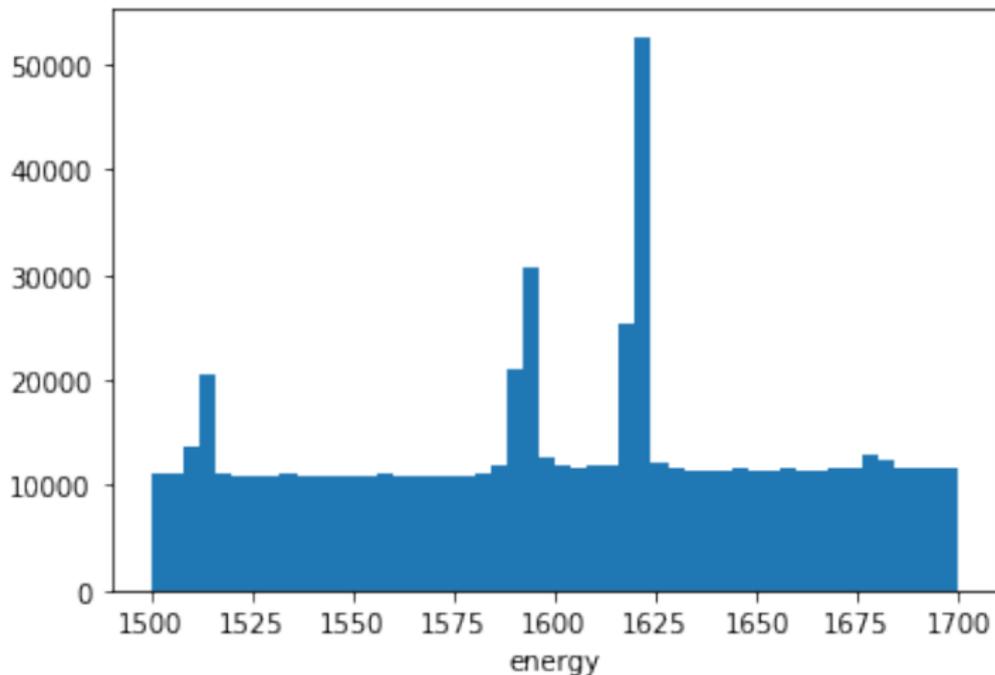
Распад ^{228}Th

^{228}Th decay chain

Energies of main α lines in keV



Распределение энергии



Гистограмма распределения энергии для калибровочных данных, используемых в анализе

Результаты

Результаты, полученные нейросетью для каждого детектора при обучении и проверке нейросети на калибровочных данных. Background rejection = 90%

Детектор	accuracy	roc_auc	f1
ANG1	0.60 ± 0.02	0.73 ± 0.02	0.42 ± 0.02
ANG2	0.66 ± 0.02	0.76 ± 0.02	0.56 ± 0.02
ANG3	0.65 ± 0.02	0.75 ± 0.02	0.53 ± 0.02
ANG4	0.66 ± 0.02	0.76 ± 0.02	0.55 ± 0.02
ANG5	0.67 ± 0.02	0.78 ± 0.02	0.58 ± 0.02
RG1	0.61 ± 0.02	0.72 ± 0.02	0.46 ± 0.02
RG2	0.64 ± 0.02	0.74 ± 0.02	0.52 ± 0.02
GTF112	0.68 ± 0.02	0.78 ± 0.02	0.58 ± 0.02
GTF45	0.65 ± 0.02	0.74 ± 0.02	0.53 ± 0.02

Результаты

Результаты, полученные нейросетью для двух детекторов при обучении и проверке нейросети на Монте-Карло данных. Background rejection = 90%

Детектор	accuracy	roc_auc	f1
ANG5	0.85 ± 0.02	0.94 ± 0.02	0.85 ± 0.02
RG2	0.93 ± 0.02	0.96 ± 0.02	0.92 ± 0.02

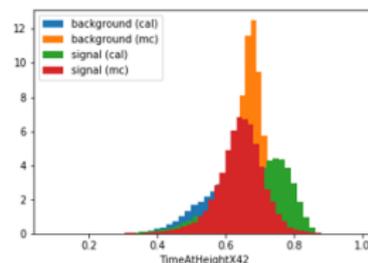
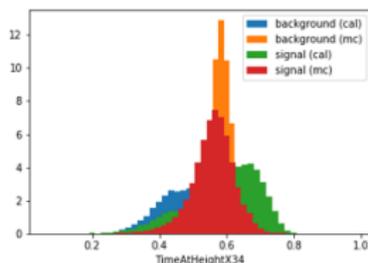
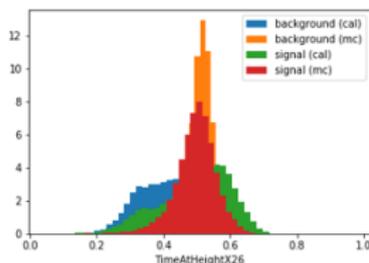
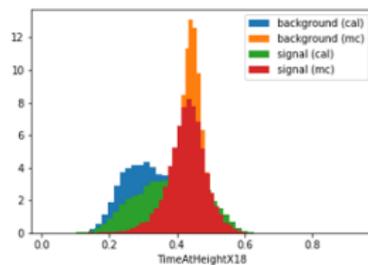
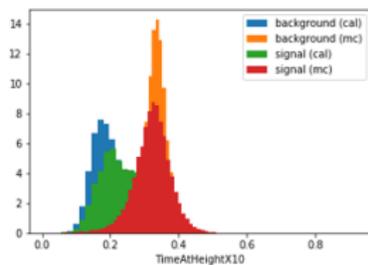
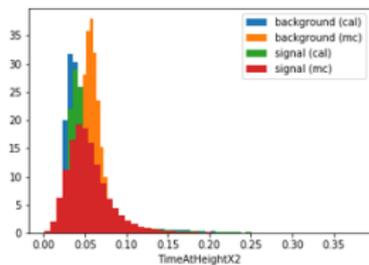
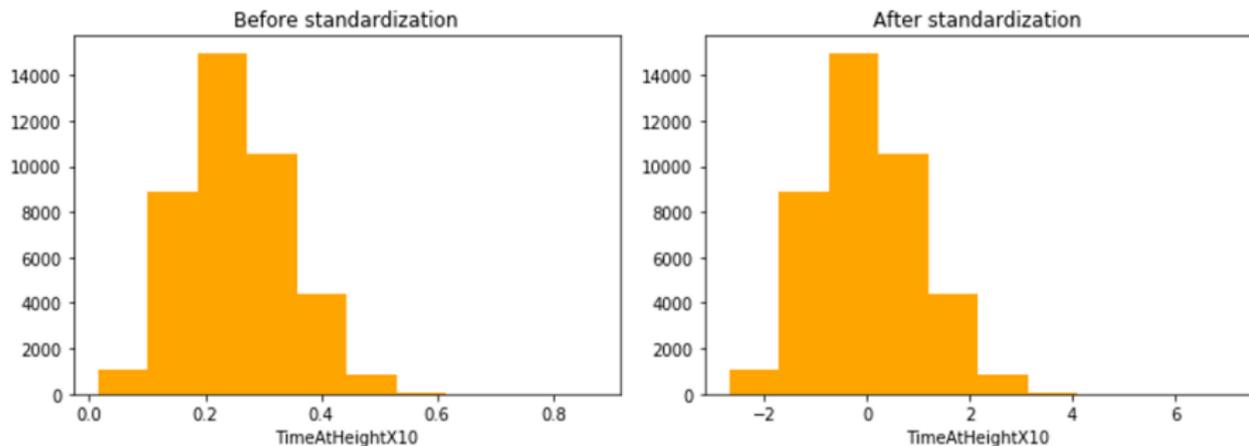
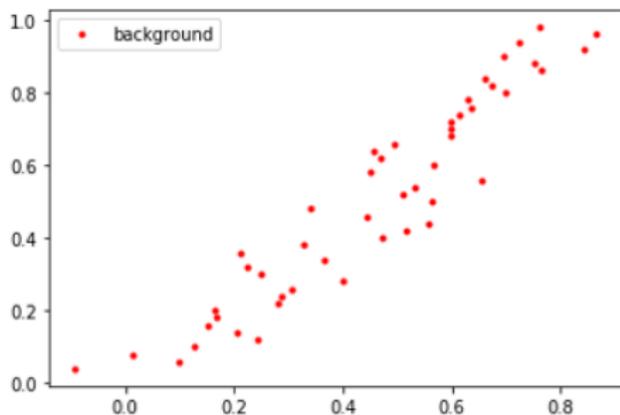
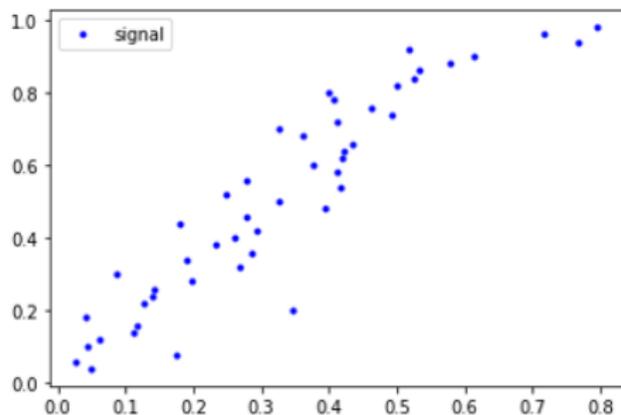


Рисунок 24 – Гистограммы некоторых параметров TimeAtHeightX в Монте-Карло данных для детектора RG2



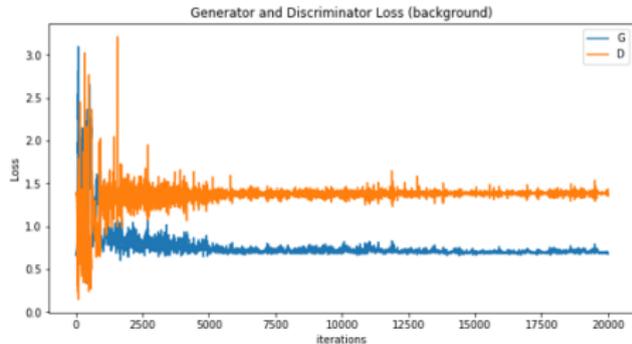
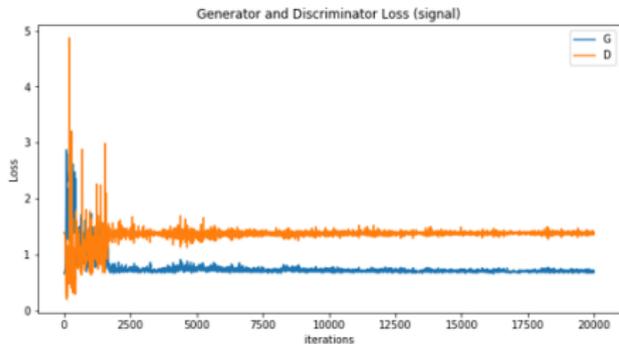
Распределения параметра TimeAtHeightX10 для сигнальных импульсов до и после стандартизации

GAN



Пример форм сигнальных и фоновых импульсов, сгенерированных перцептроном, обученным на нестандартизованных данных

GAN



Функции потерь при обучении моделей генератора и дискриминатора, состоящих из линейных слоев, для сигнала и фона

Результаты, полученные нейросетью для детектора ANG5 при обучении нейросети на сгенерированных данных и ее проверке на физических (калибровочных) данных

Детектор	accuracy	roc_auc	f1
ANG5	0.60 ± 0.02	0.61 ± 0.02	0.58 ± 0.02