Моделирование формы импульса при помощи генеративно-состязательных нейронных сетей для полукоаксиальных детекторов в эксперименте GERDA

Зарецкий Н.Д.
Научный руководитель: Черногоров А.Е.

Национальный исследовательский ядерный университет «МИФИ»

22 декабря 2021 г.

Введение

Цель работы: моделирование формы импульса для полукоаксиальных детекторов в эксперименте GERDA (phase II) с помощью генеративно-состязательных нейронных сетей и использование синтезированных импульсов для задачи разделения сигнальных и фоновых событий.

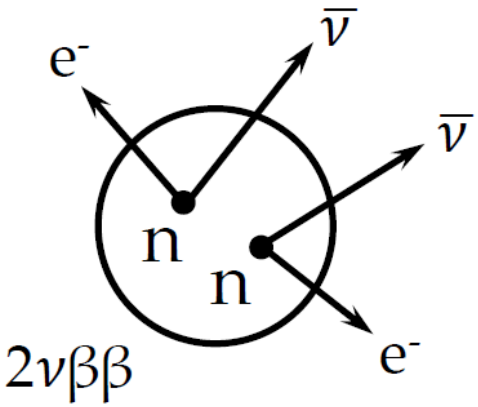


Рисунок 1 – Схема $2\nu\beta\beta$

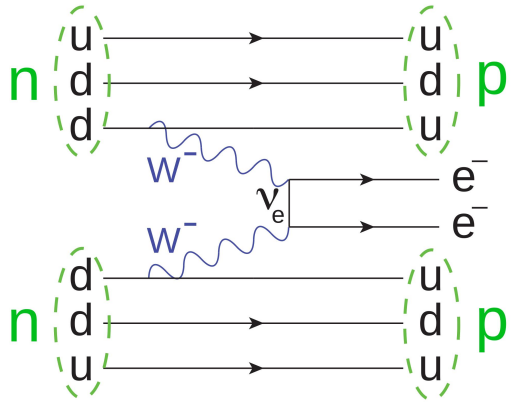


Рисунок 2 – Диаграмма $0\nu\beta\beta$

Эксперимент GERDA

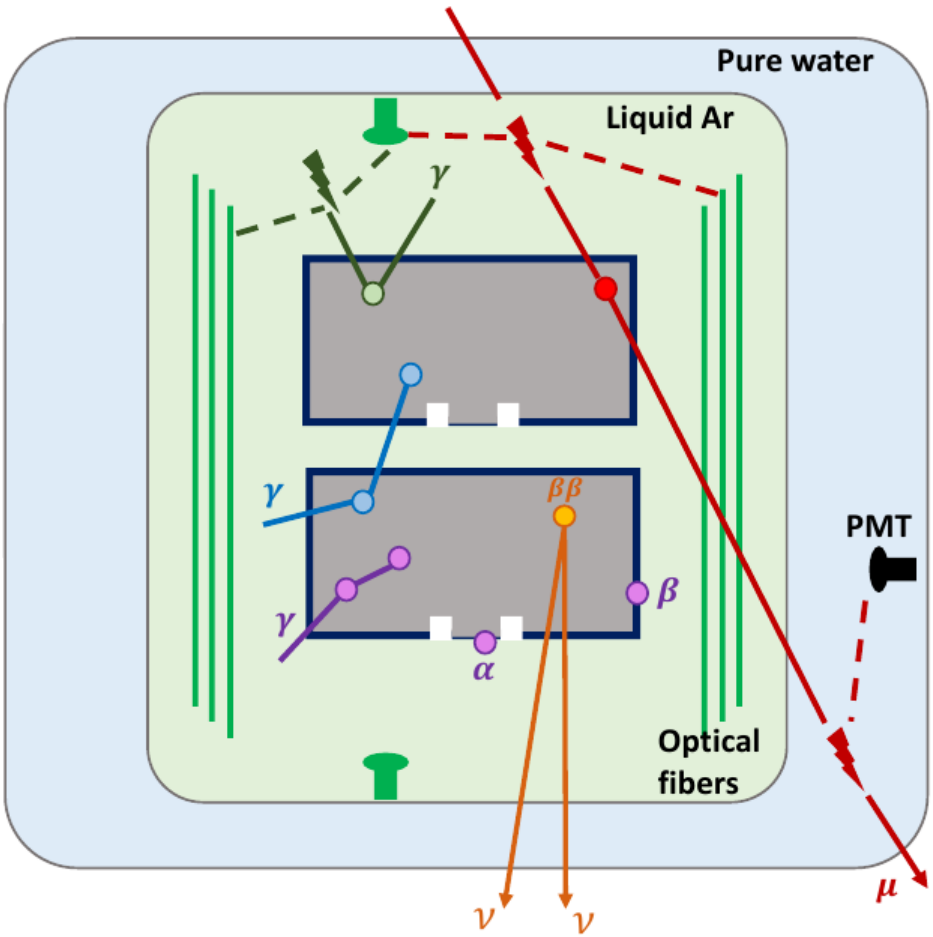


Рисунок 3 – Схема эксперимента GERDA

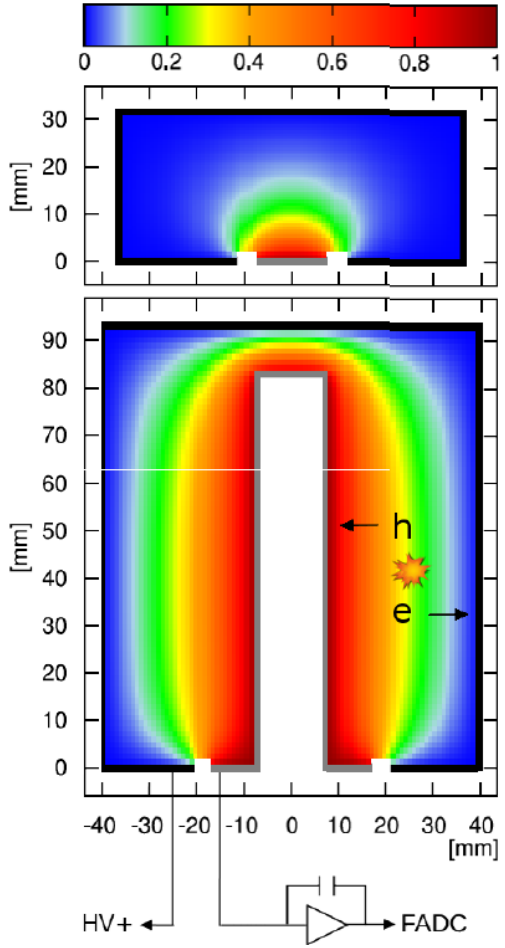


Рисунок 4 – Схема полукоаксиальных и BEGe детекторов

Анализ калибровочных данных

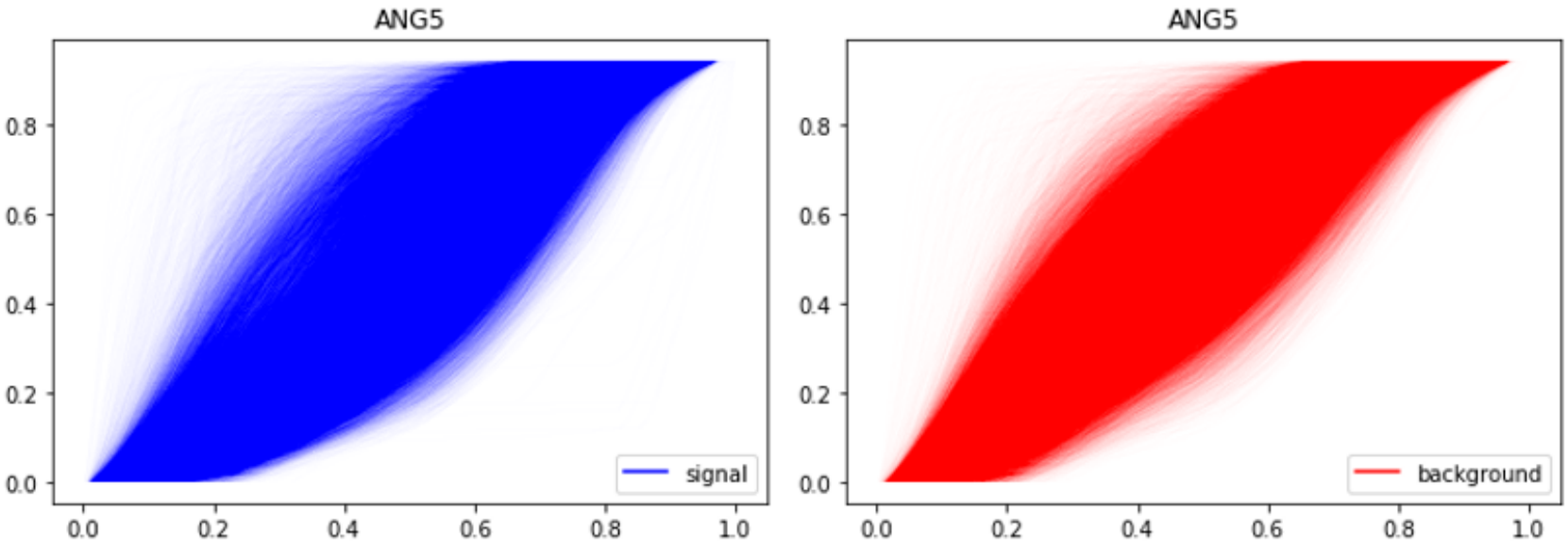


Рисунок 5 – Формы сигнальных и фоновых импульсов, полученные в эксперименте GERDA. По горизонтальной оси - время, нормированное на длительность импульса

Анализ калибровочных данных

Таблица 1 – Результаты, полученные для детектора ANG5 при обучении и проверке классификаторов на калибровочных данных

	accuracy	roc_auc	f1
Decision Tree	0.64 ± 0.02	0.64 ± 0.02	0.52 ± 0.02
Random Forest	0.68 ± 0.02	0.68 ± 0.02	0.68 ± 0.02

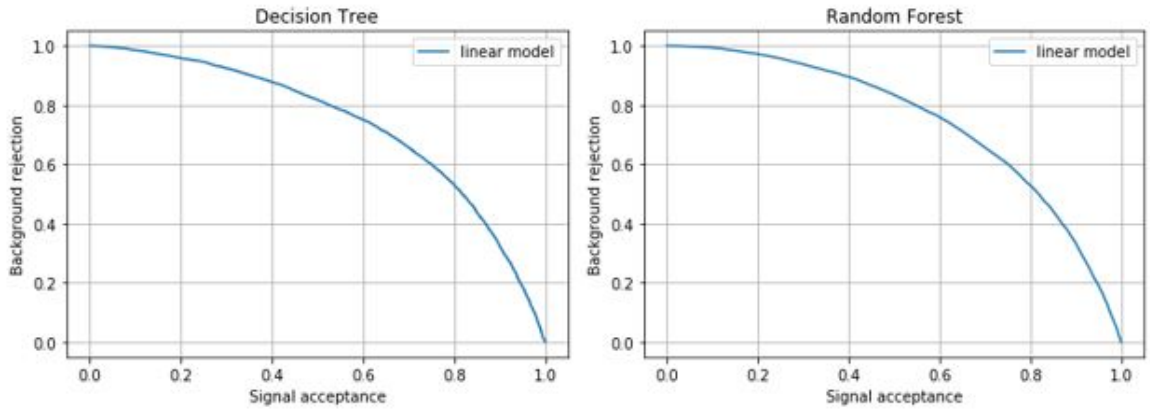


Рисунок 6 – ROC-кривые для детектора ANG5 при обучении и проверке классификаторов на калибровочных данных. Background rejection - TNR (1-FPR), signal efficiency - TPR

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

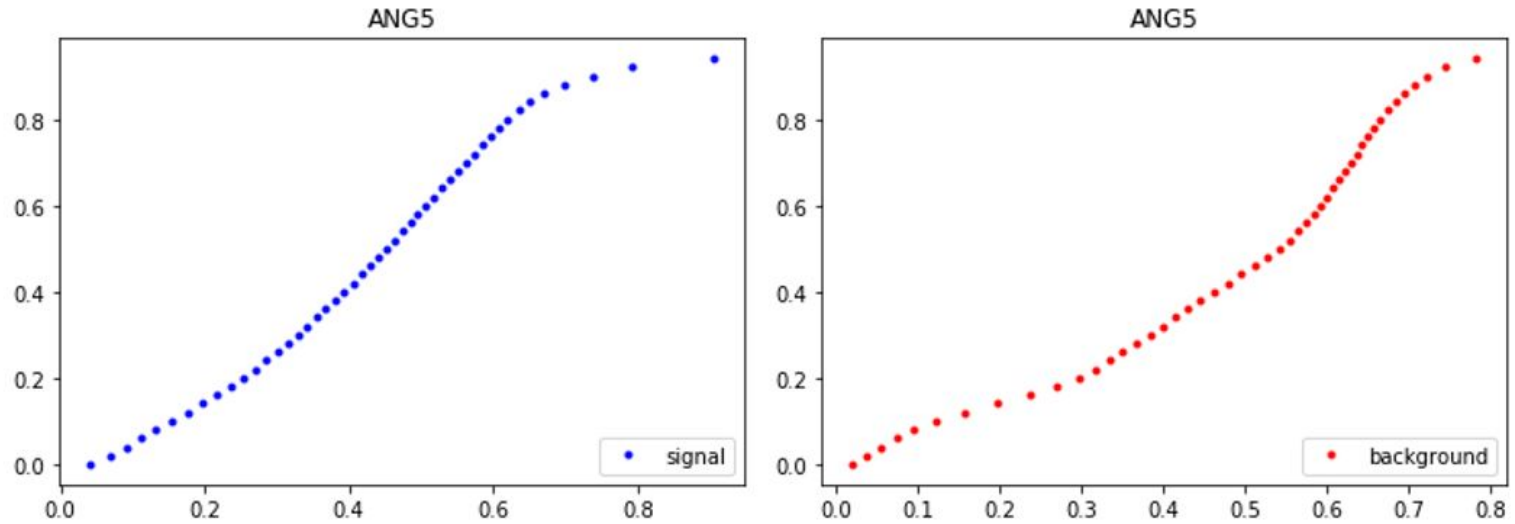


Рисунок 7 – Пример форм сигнальных и фоновых импульсов, полученных в эксперименте GERDA. По горизонтальной оси - время, нормированное на длительность импульса

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

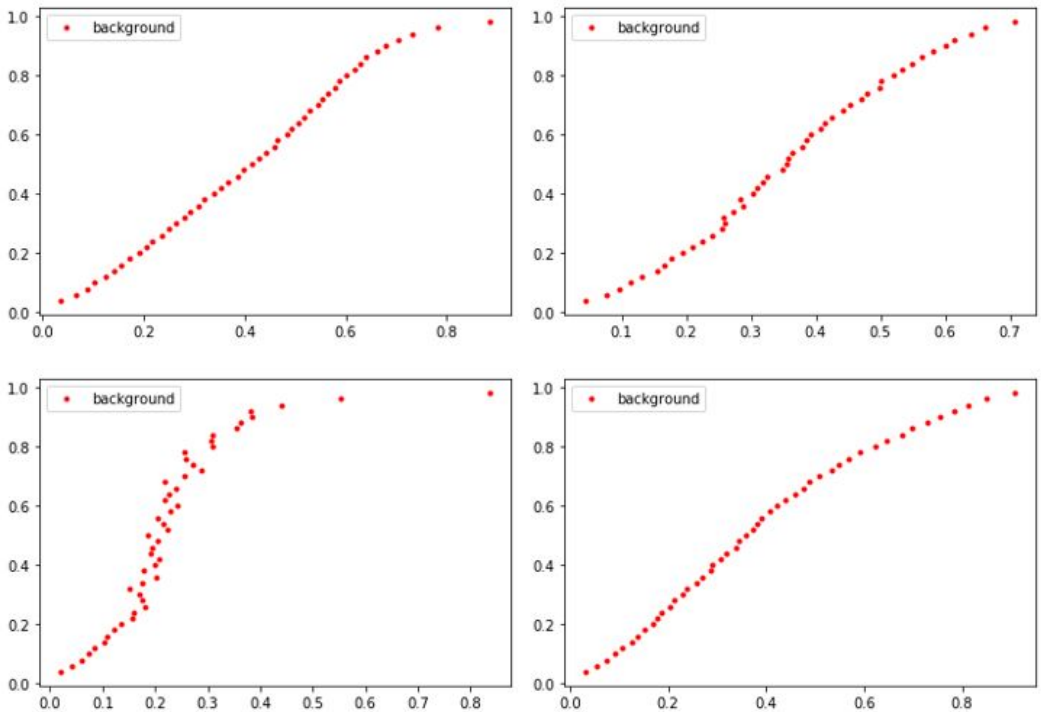


Рисунок 8 – Примеры синтезированных импульсов для детектора ANG5 (Wasserstein GAN). По горизонтальной оси - время, нормированное на длительность импульса

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

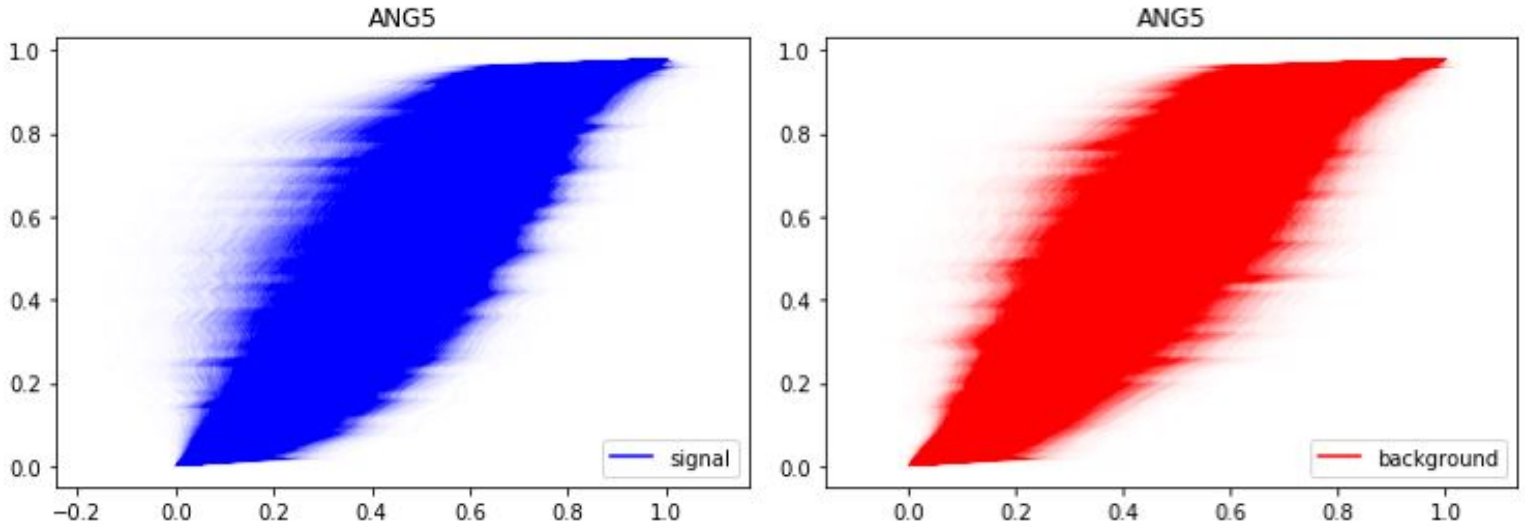


Рисунок 9 – Формы синтезированных импульсов для сигнала и фона для детектора ANG5 (Wasserstein GAN, 60000 событий). По горизонтальной оси - время, нормированное на длительность импульса

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

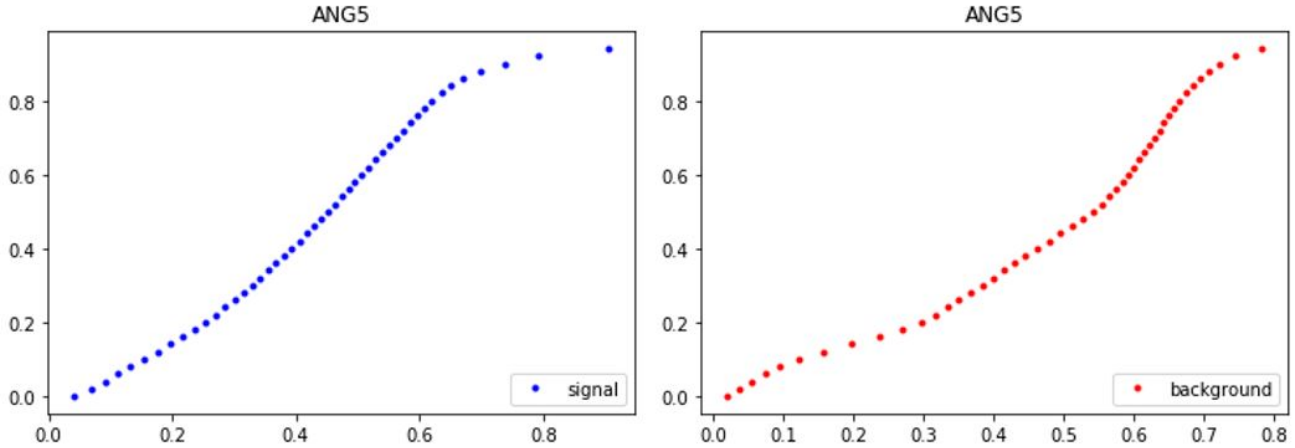


Рисунок 10 – Пример форм сигнальных и фоновых импульсов, полученных в эксперименте GERDA. По горизонтальной оси - время, нормированное на длительность импульса

$$current[i] = \frac{0.02}{TimeAtHeightX[i+1] - TimeAtHeightX[i]}$$

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

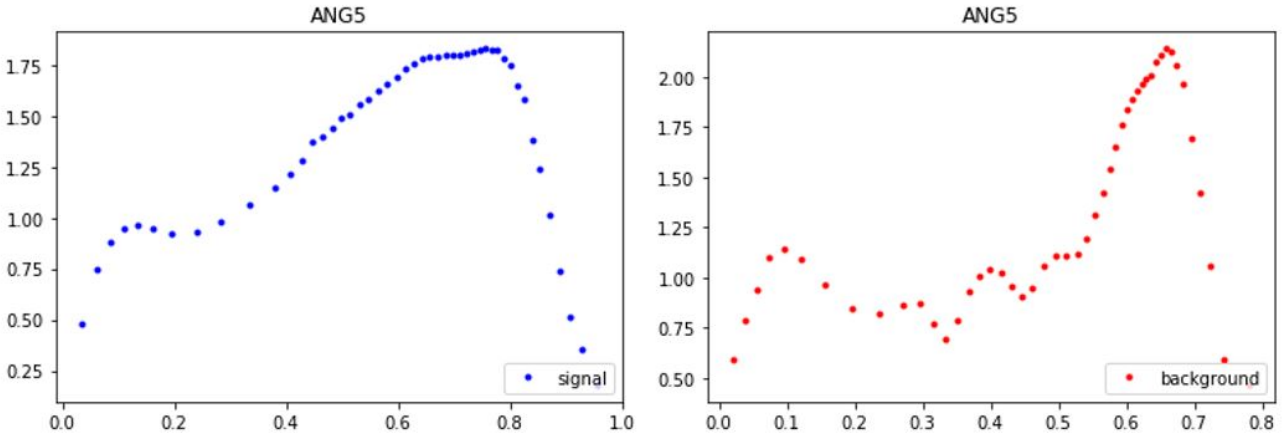


Рисунок 11 – Пример форм сигнальных и фоновых импульсов, полученных в эксперименте GERDA. По горизонтальной оси - время, нормированное на длительность импульса, по вертикальной оси - значения, полученные по формуле (1)

$$current[i] = \frac{0.02}{TimeAtHeightX[i+1] - TimeAtHeightX[i]} \quad (1)$$

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

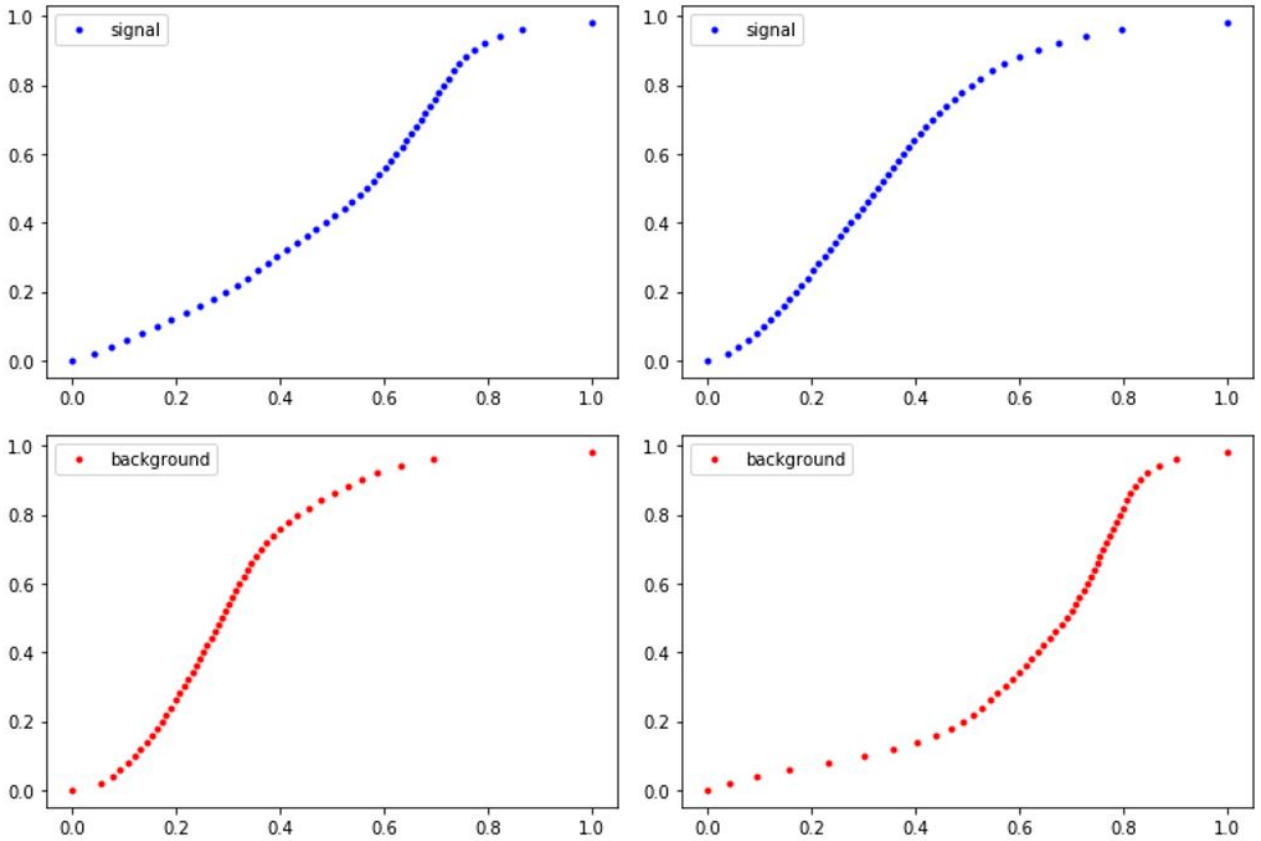


Рисунок 12 – Примеры синтезированных импульсов для детектора ANG5 (Wasserstein GAN). По горизонтальной оси - время, нормированное на длительность импульса

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

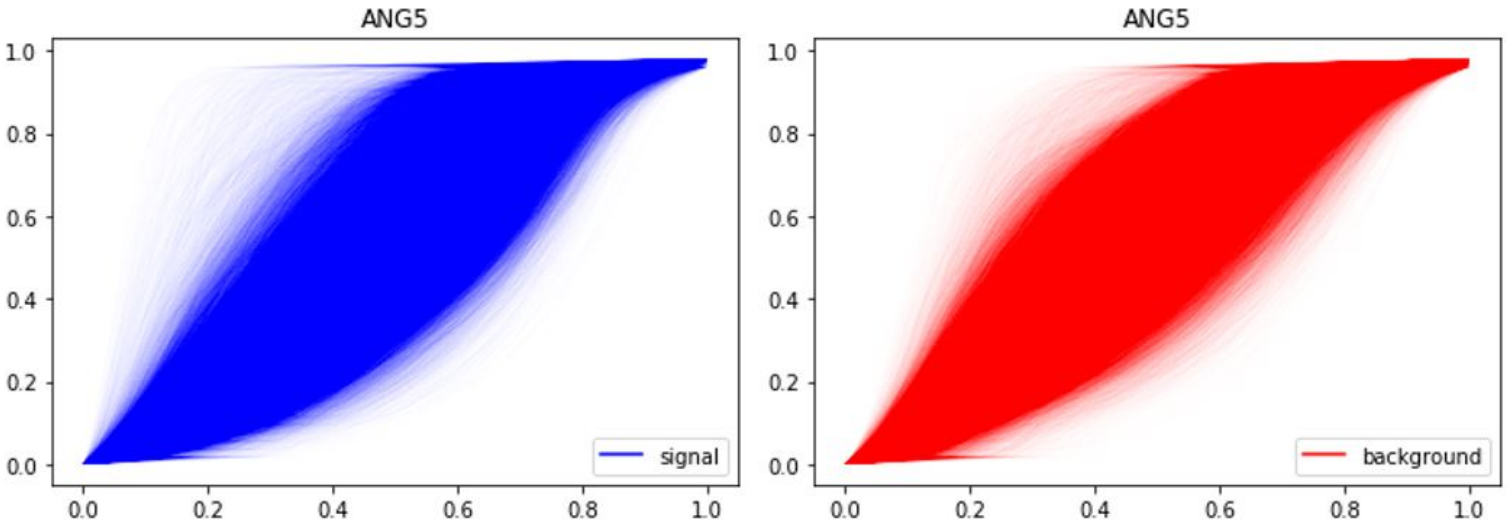


Рисунок 13 – Примеры синтезированных импульсов для детектора ANG5 (Wasserstein GAN). По горизонтальной оси - время, нормированное на длительность импульса

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

Таблица 2 – Результаты, полученные для детектора ANG5 при обучении классификаторов на синтетических данных (Wasserstein GAN) и их проверке на калибровочных данных

	accuracy	roc_auc	f1
Decision Tree	0.65 ± 0.02	0.65 ± 0.02	0.64 ± 0.02
Random Forest	0.65 ± 0.02	0.65 ± 0.02	0.63 ± 0.02

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

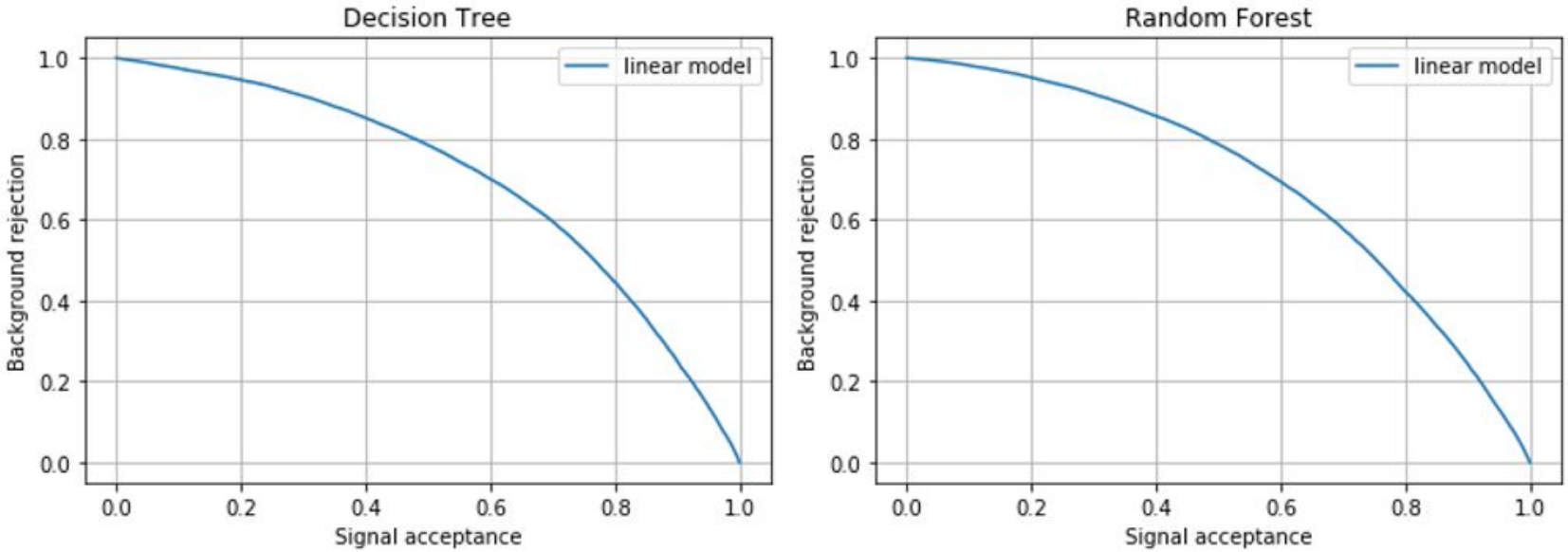


Рисунок 14 – ROC-кривые для детектора ANG5 при обучении классификаторов на синтетических данных (Wasserstein GAN) и их проверке на калибровочных данных. Background rejection - TNR (1-FPR), signal efficiency - TPR

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

Таблица 3 – Результаты, полученные для детектора ANG5 при разделении синтетических (Wasserstein GAN) и реальных данных

	accuracy	roc_auc	f1
Decision Tree	0.64 ± 0.02	0.64 ± 0.02	0.60 ± 0.02
Random Forest	0.69 ± 0.02	0.69 ± 0.02	0.69 ± 0.02

Моделирование сигнальных и фоновых событий с помощью WGAN-сетей

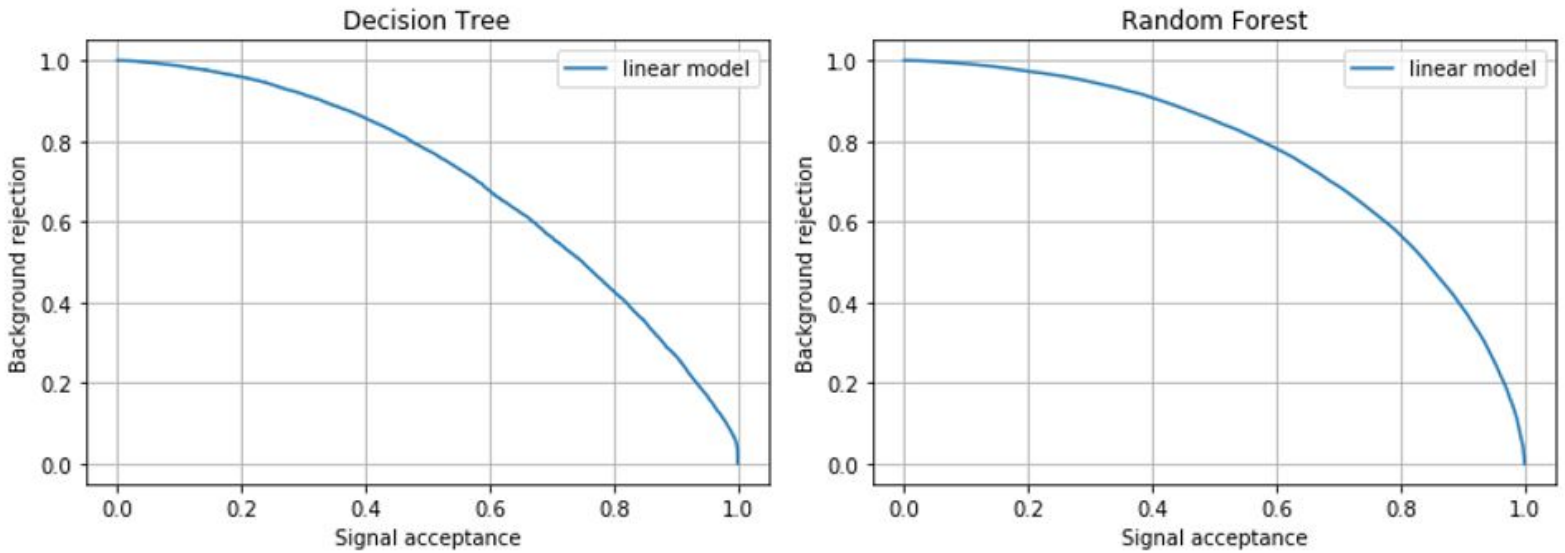


Рисунок 15 – ROC-кривые для детектора ANG5 при разделении синтетических (Wasserstein GAN) и реальных данных. Background rejection - TNR (1-FPR), signal efficiency - TPR

Заключение

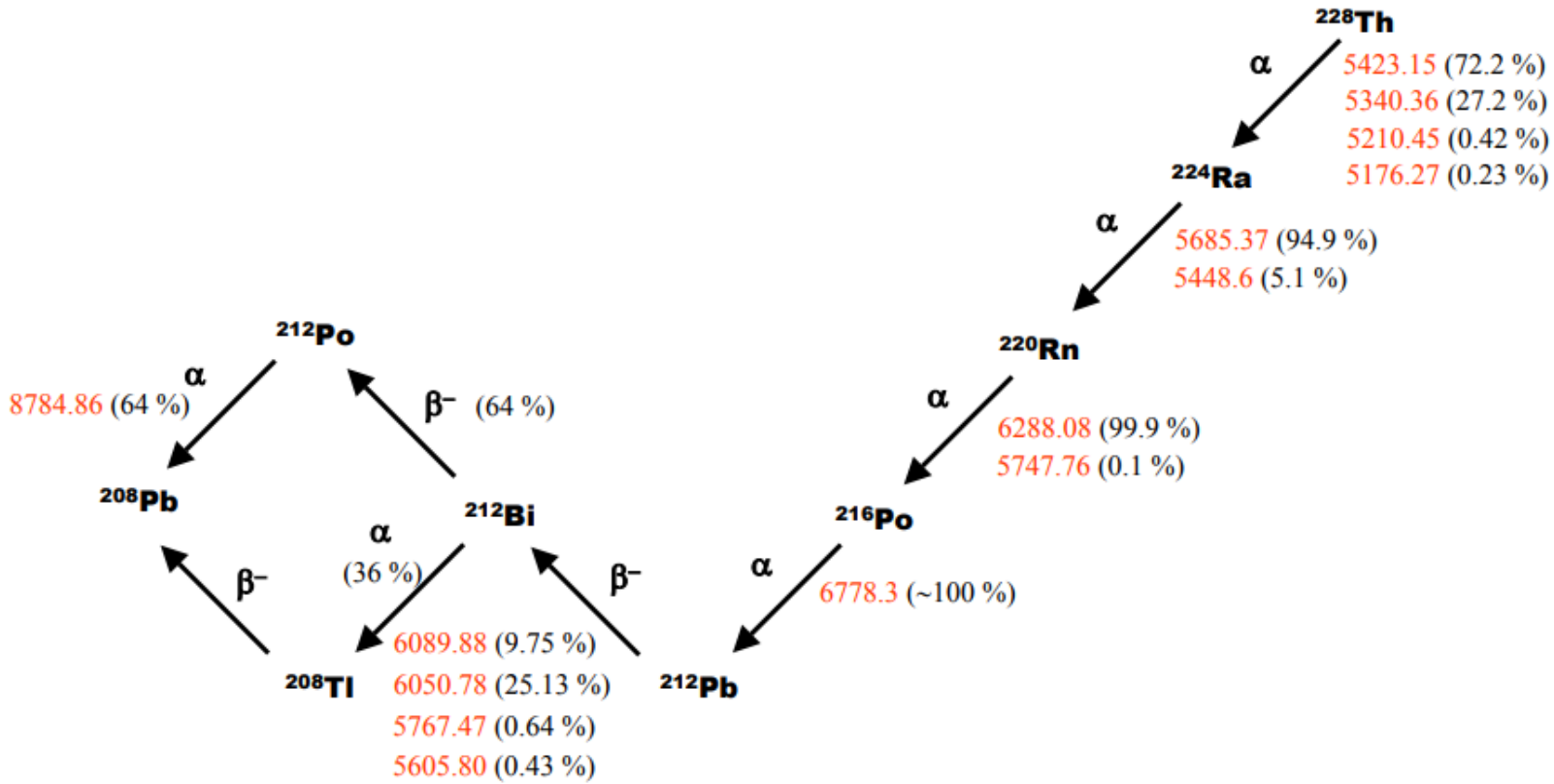
В ходе работы были сделаны следующие выводы:

- результаты обучения классификаторов на синтезированных с помощью GAN сетей данных соответствуют эталонным данным.
- при разделении синтезированных и реальных данных классификаторами результат выше (то есть хуже) ожидаемого.

Таким образом, на данном этапе невозможно однозначно утверждать, что созданные GAN-сети в процессе обучения выдают импульсы, соответсвующие реальным данным.
В дальнейшем планируется распространить приведенный анализ на другие типы детекторов.

Дополнительные слайды

Распад ^{228}Th



Калибровки

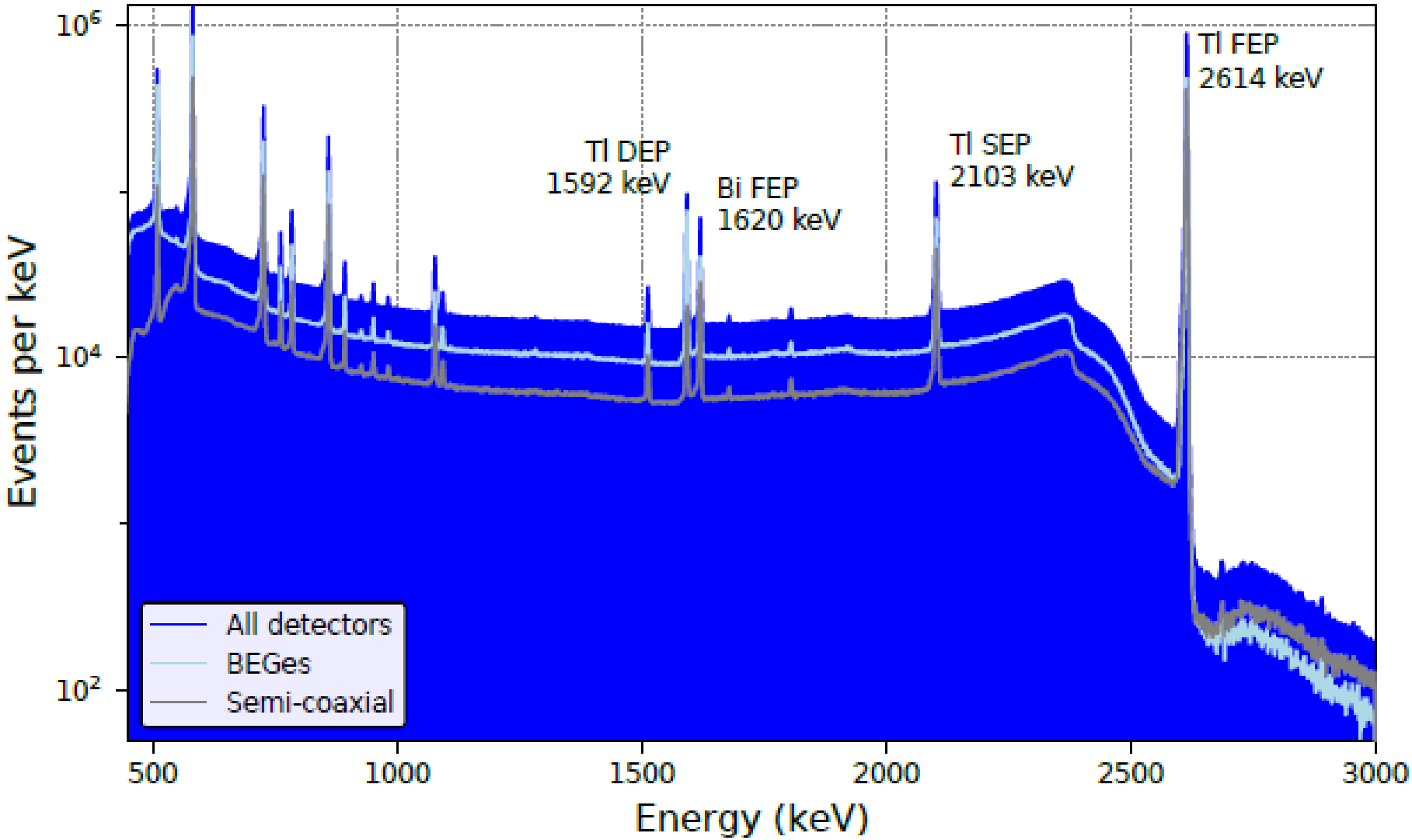
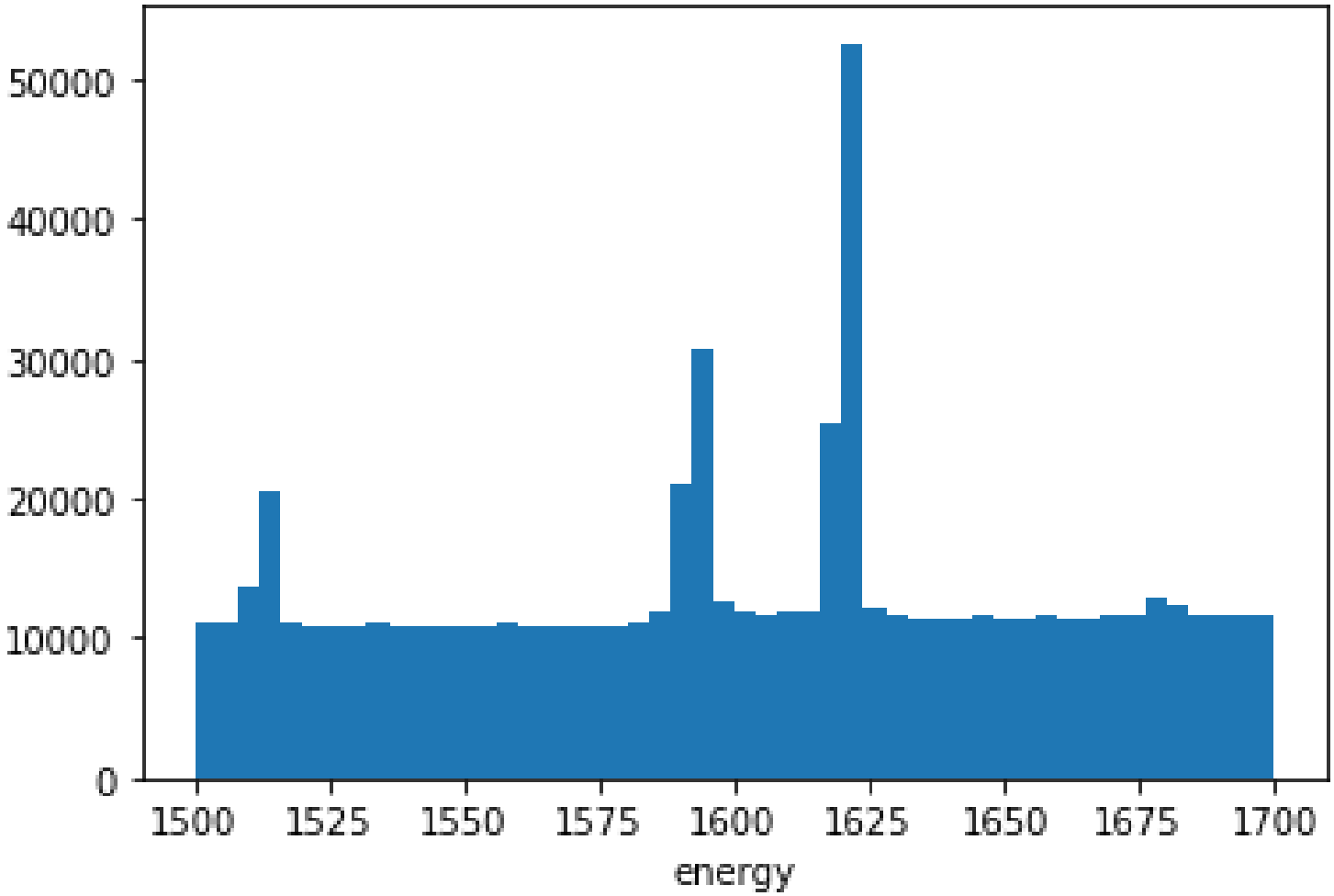


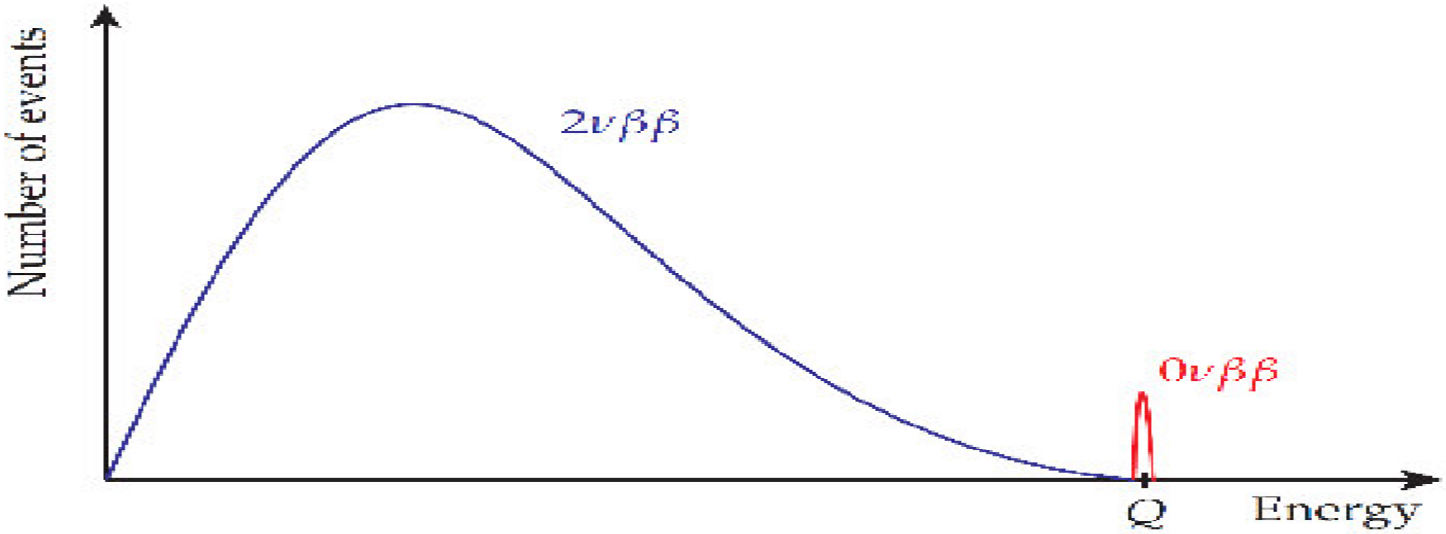
Рисунок 16 – Энергетический спектр $^{228}\mathrm{Th}$

Распределение энергии

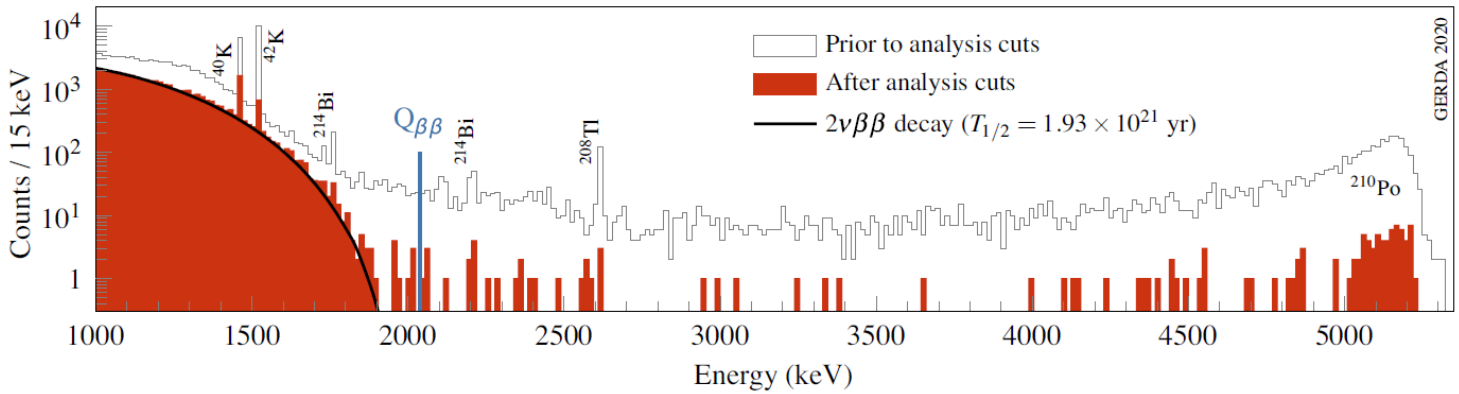


Гистограмма распределения энергии для калибровочных данных, используемых в анализе

Данные с эксперимента GERDA

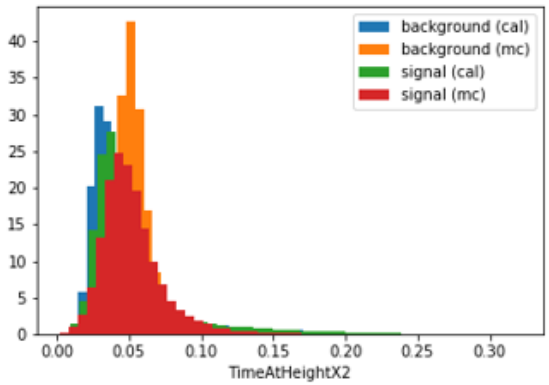


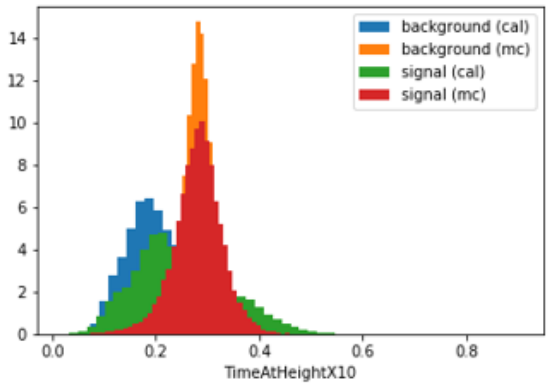
Энергетический спектр двух электронов

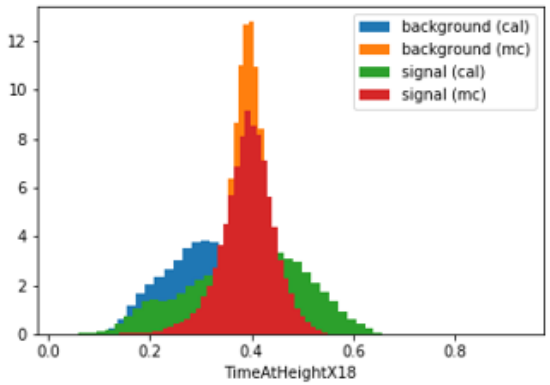


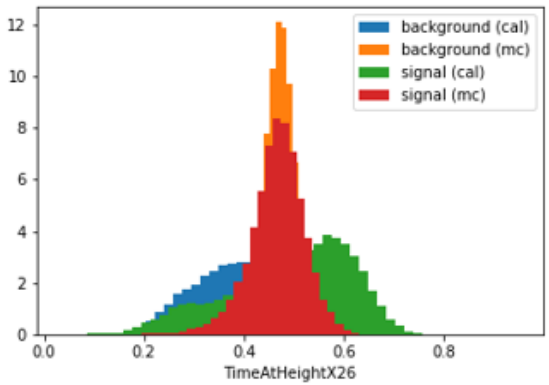
Спектр энергии с эксперимента GERDA

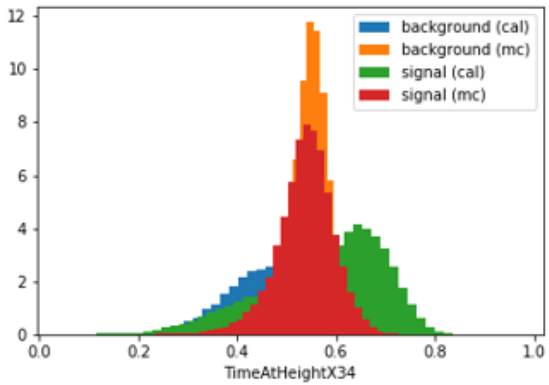
Монте-Карло данные

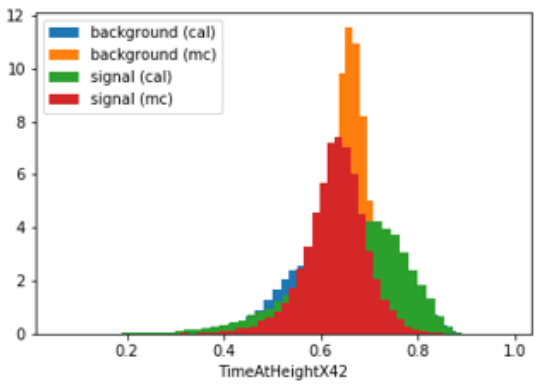






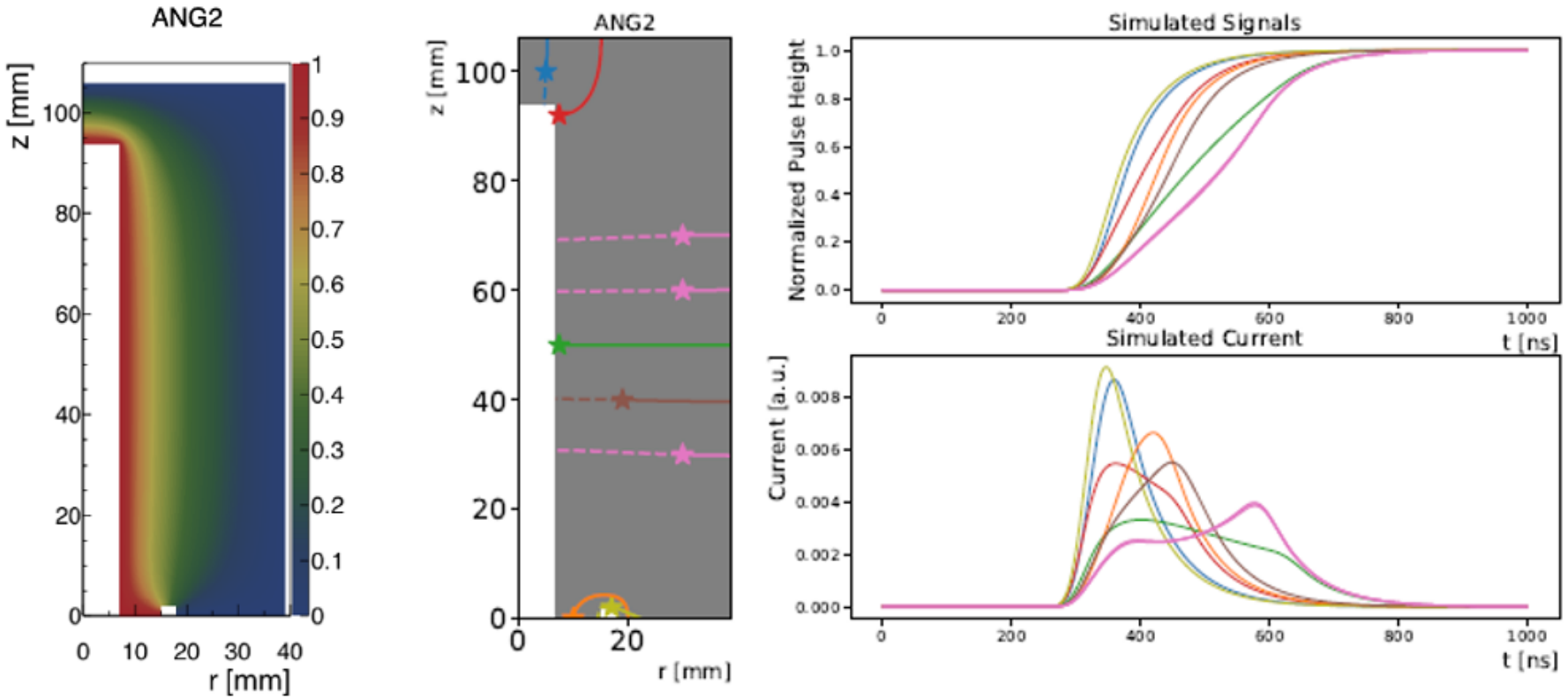






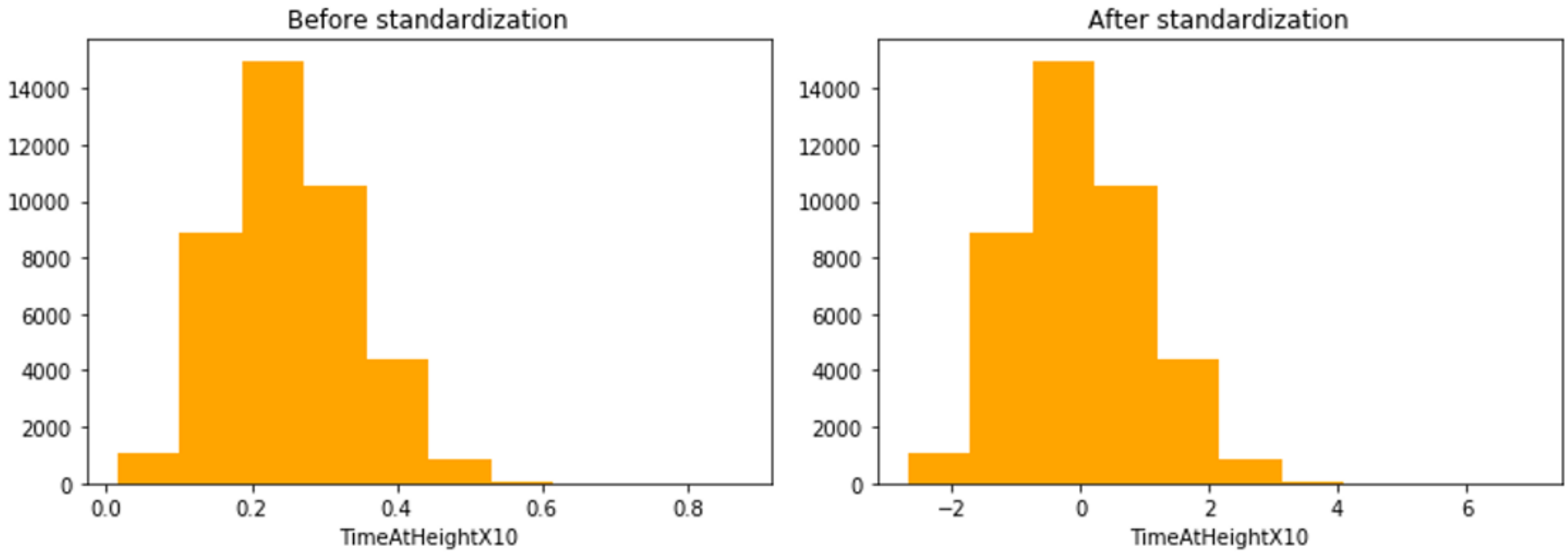
Гистограммы некоторых параметров TimeAtHeightX в Монте-Карло данных для детектора ANG5

Формы импульсов для полукоаксиальных детекторов



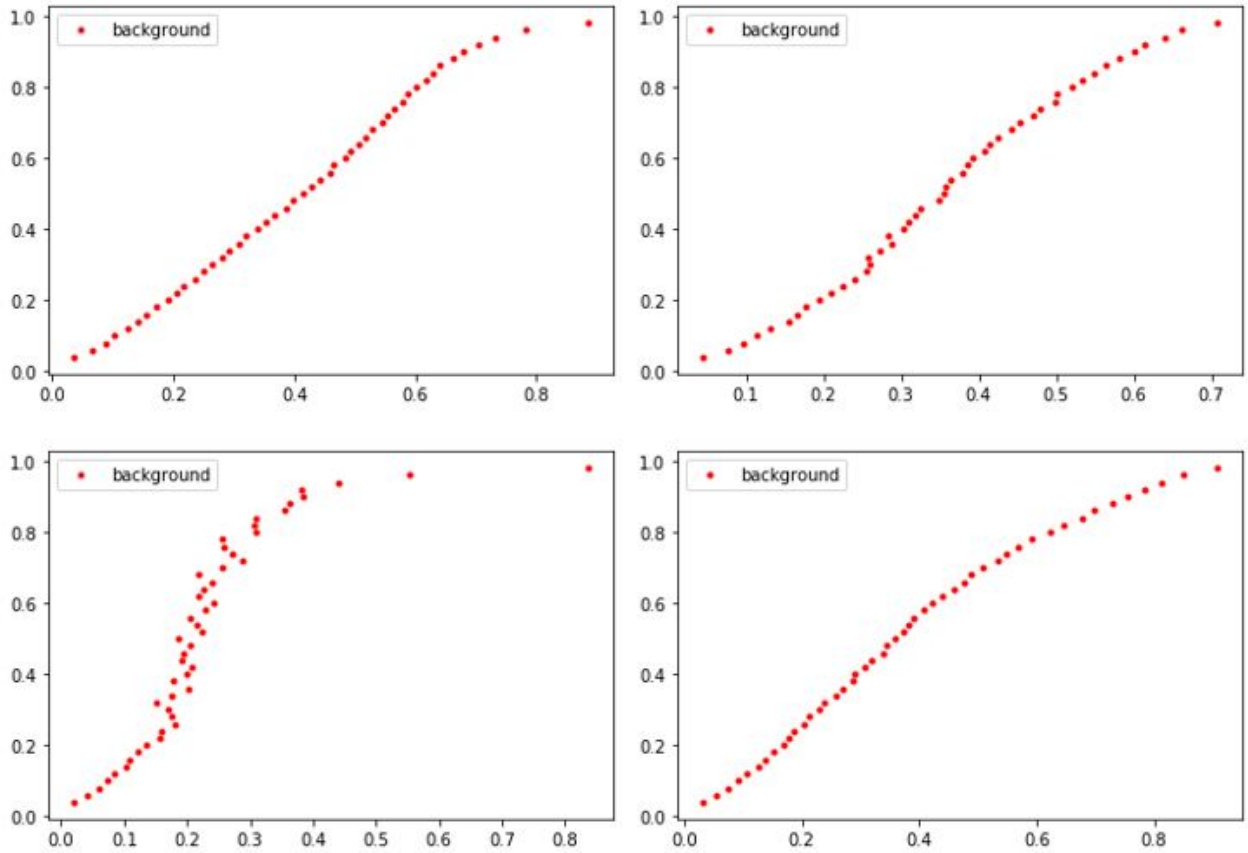
Возможные формы импульсов для разных мест выделения энергии в полукоаксиальном детекторе

GAN



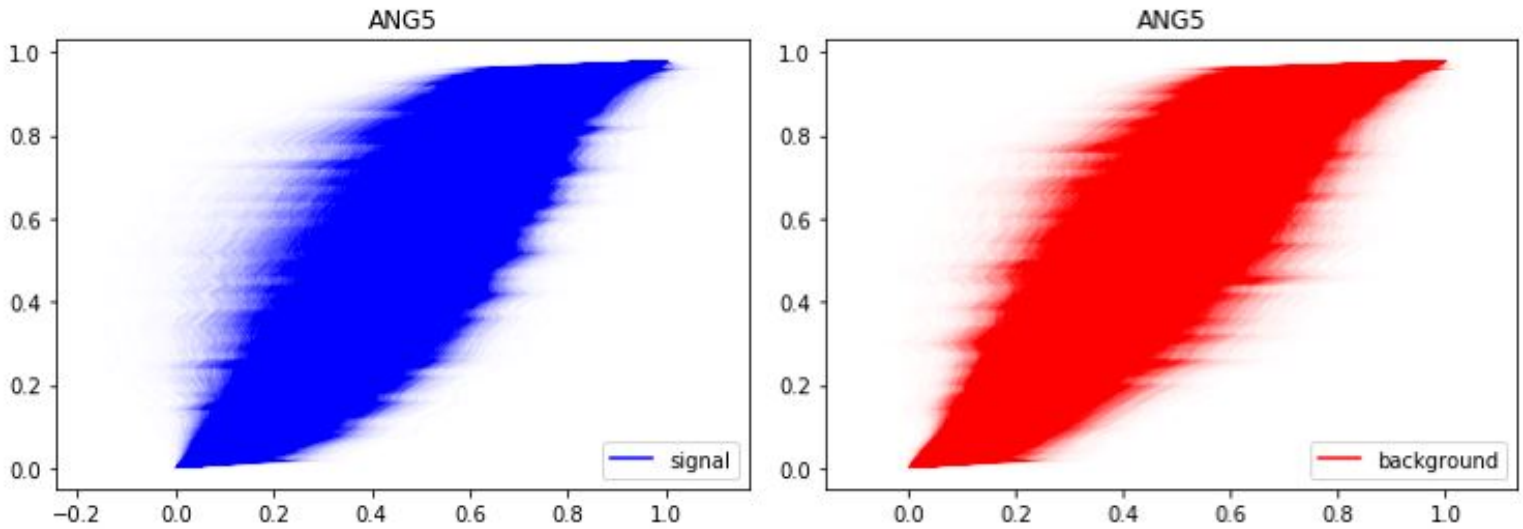
Распределения параметра TimeAtHeightX10 для сигнальных импульсов до и после стандартизации

Minimax GAN, входные данные - импульсы заряда



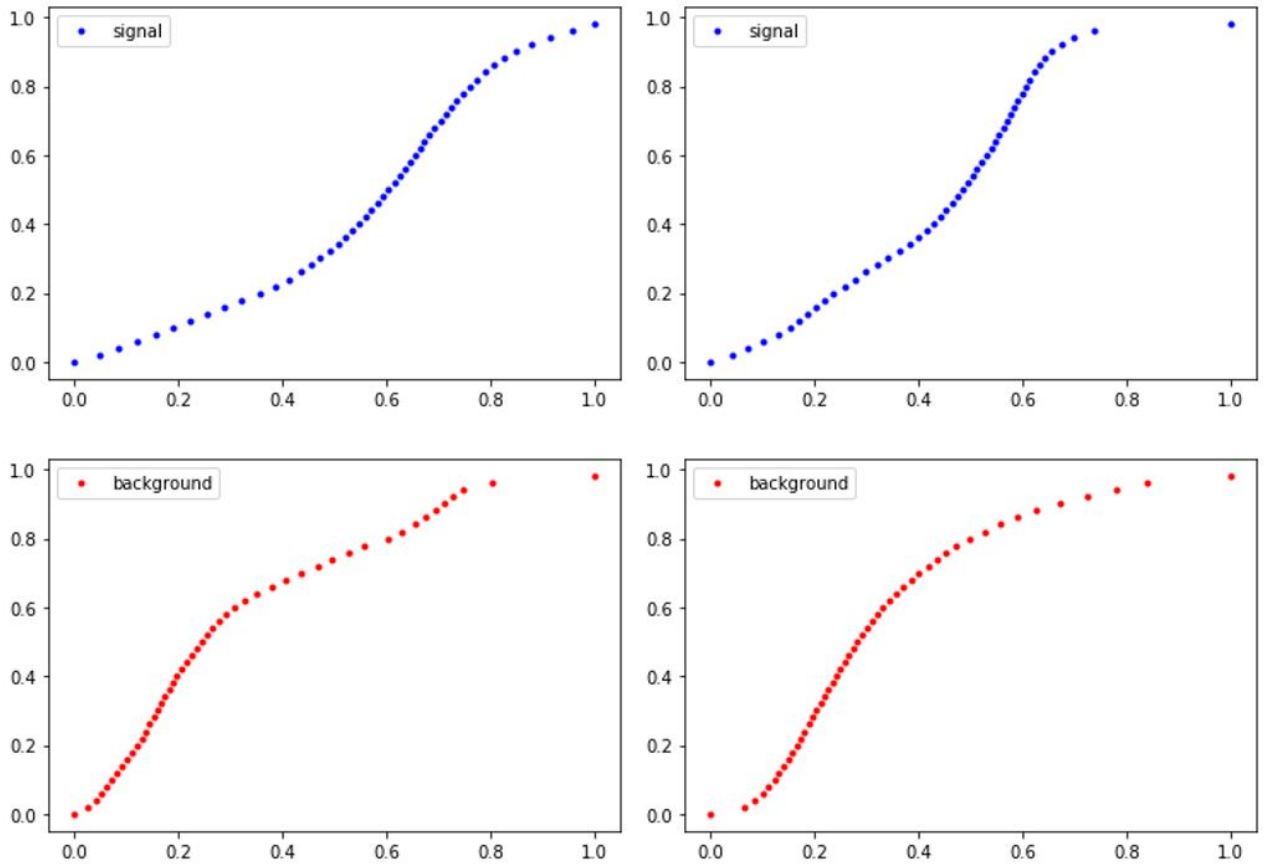
Примеры синтезированных импульсов для детектора ANG5 (Minimax GAN). По горизонтальной оси - время, нормированное на длительность импульса

Minimax GAN, входные данные - импульсы заряда



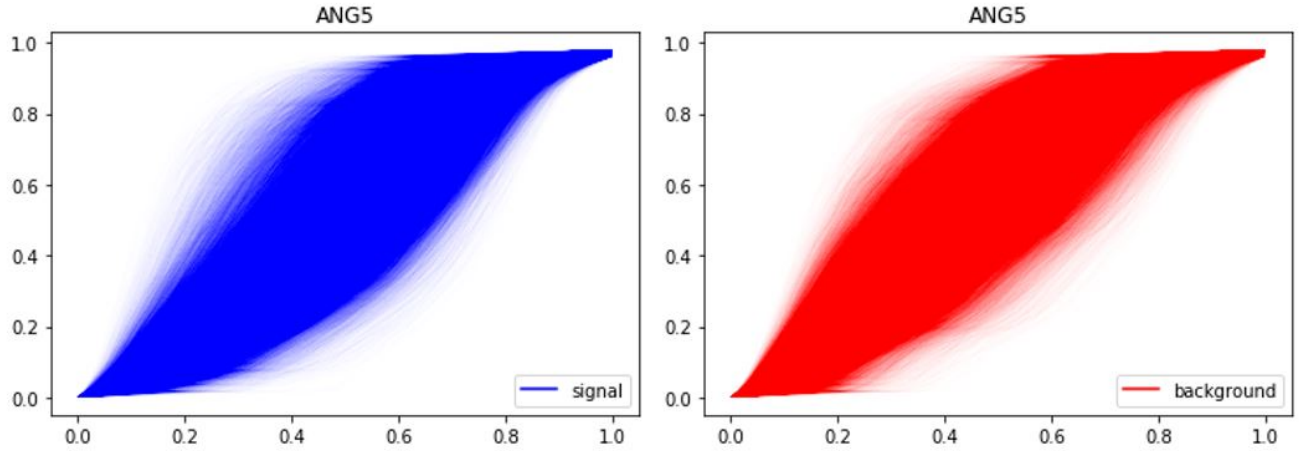
Формы синтезированных импульсов для сигнала и фона для детектора ANG5 (Minimax GAN, 30000 событий). По горизонтальной оси - время, нормированное на длительность импульса

Minimax GAN, входные данные - импульсы заряда

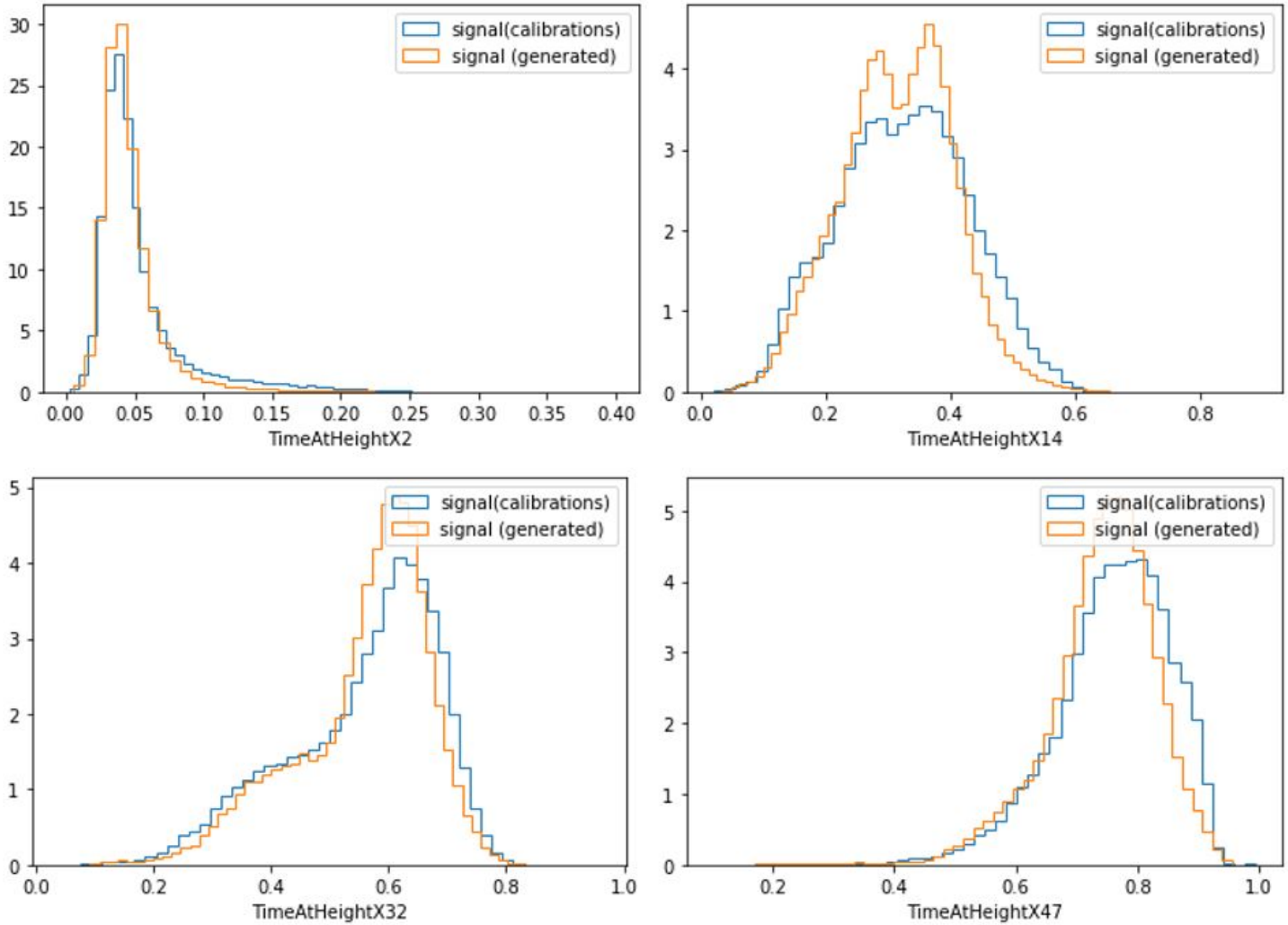


Примеры синтезированных импульсов для детектора ANG5 (Minimax GAN). По горизонтальной оси - время, нормированное на длительность импульса

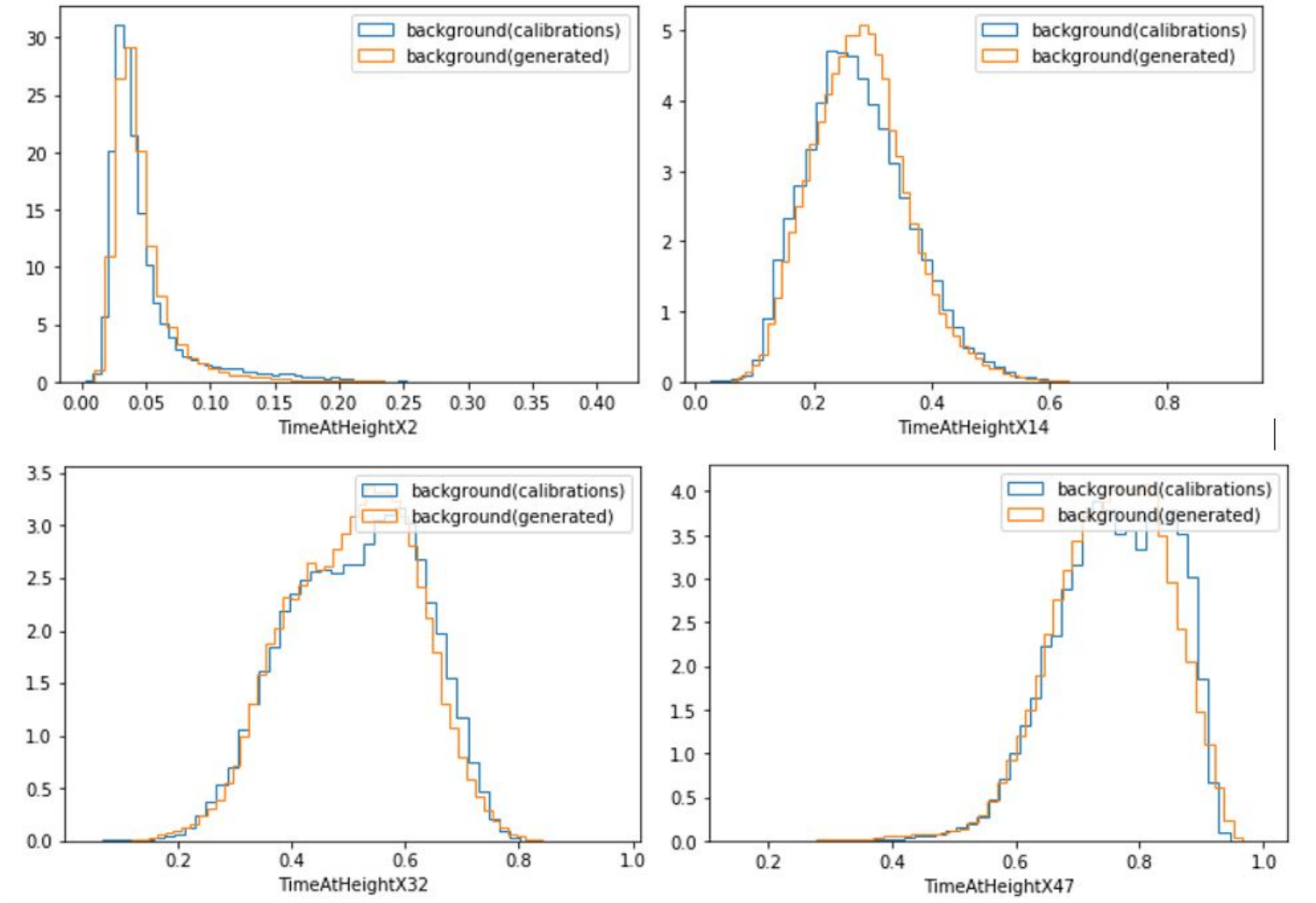
Minimax GAN, входные данные - импульсы заряда



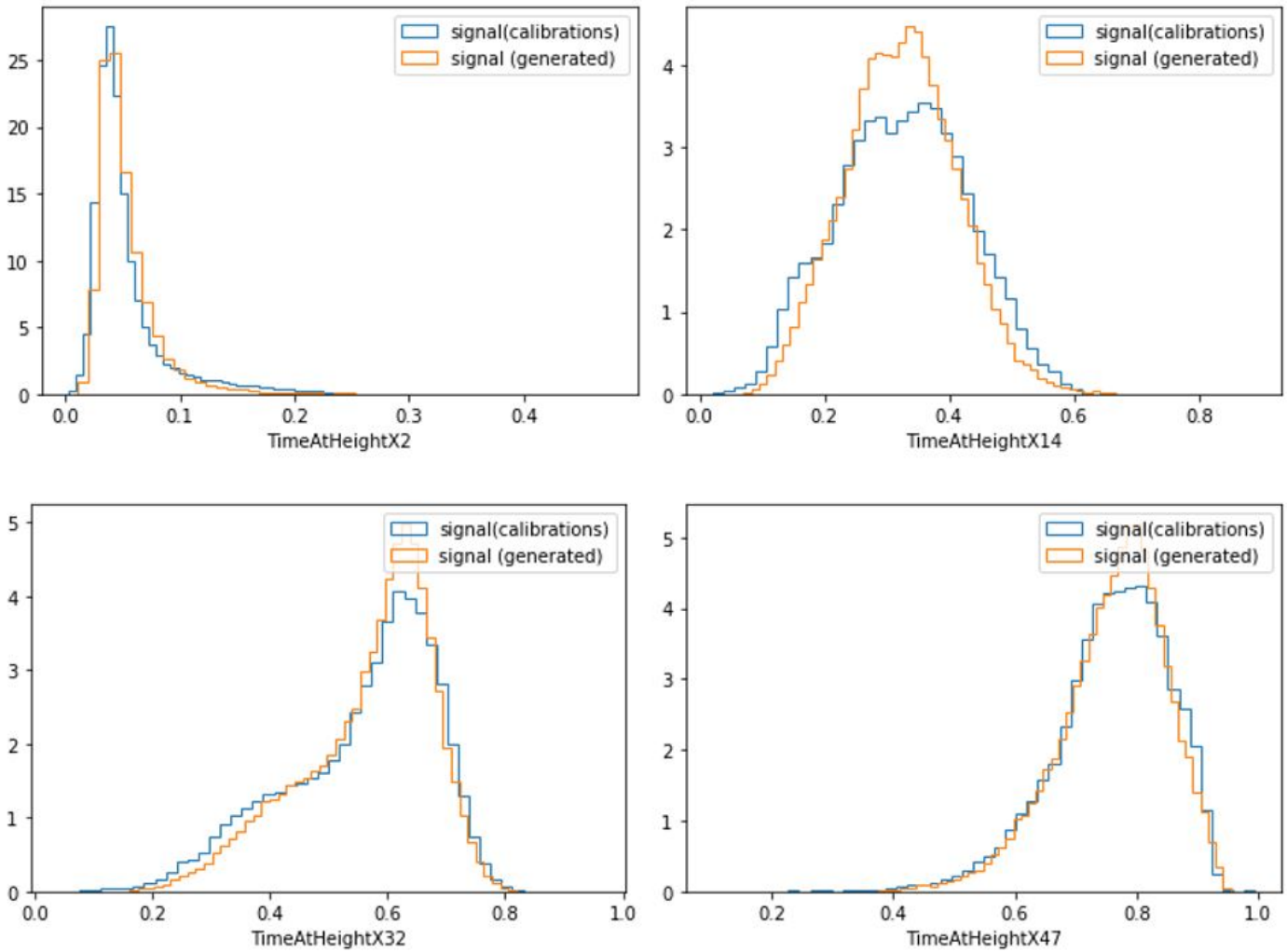
Примеры синтезированных импульсов для детектора ANG5 (Minimax GAN). По горизонтальной оси - время, нормированное на длительность импульса



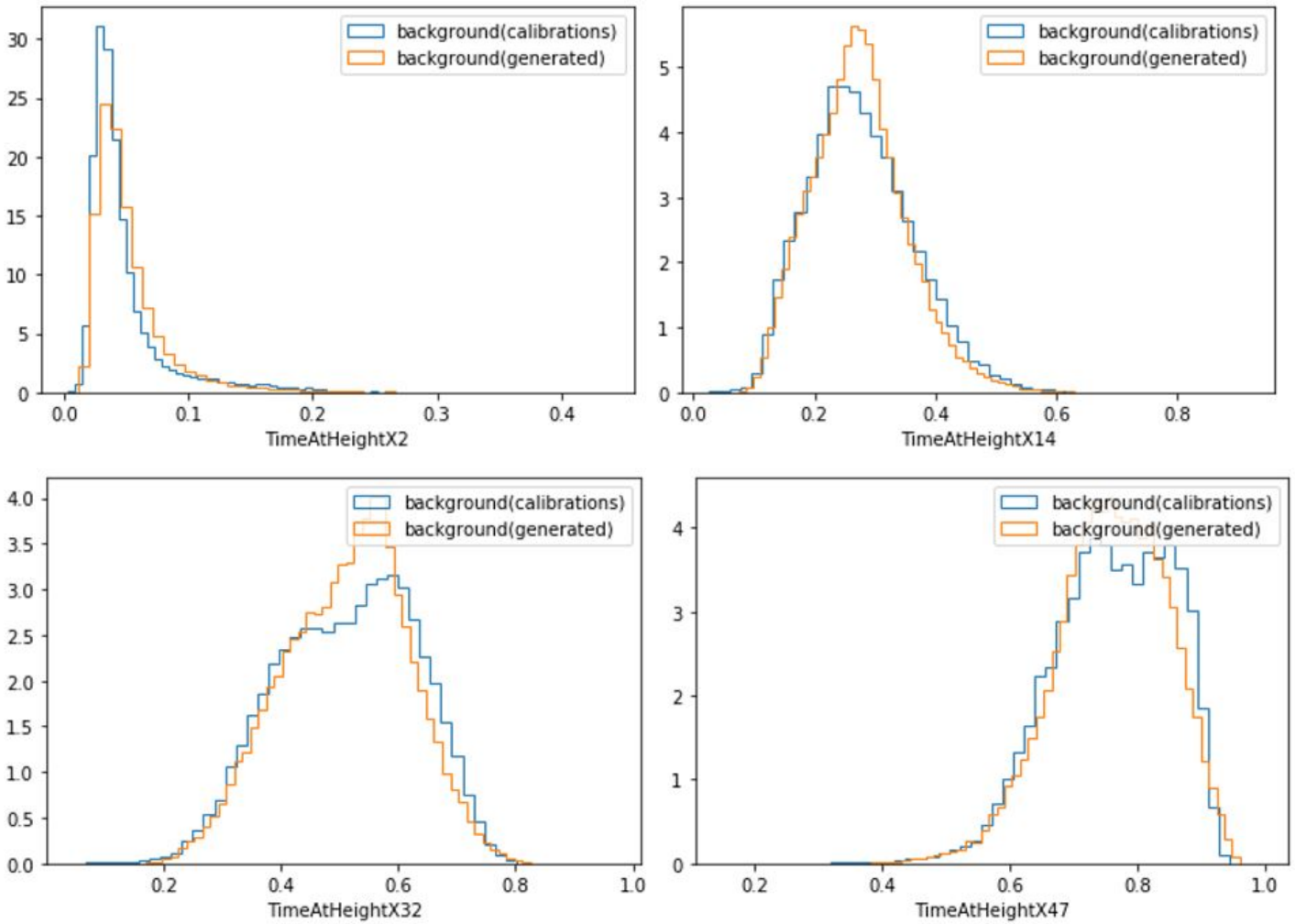
Сравнение распределений некоторых параметров TimeAtHeightX для синтезированных (Wasserstein GAN) и калибровочных сигнальных событий для детектора ANG5



Сравнение распределений некоторых параметров TimeAtHeightX для синтезированных (Wasserstein GAN) и калибровочных фоновых событий для детектора ANG5



Сравнение распределений некоторых параметров TimeAtHeightX для синтезированных (Minimax GAN) и калибровочных сигнальных событий для детектора ANG5



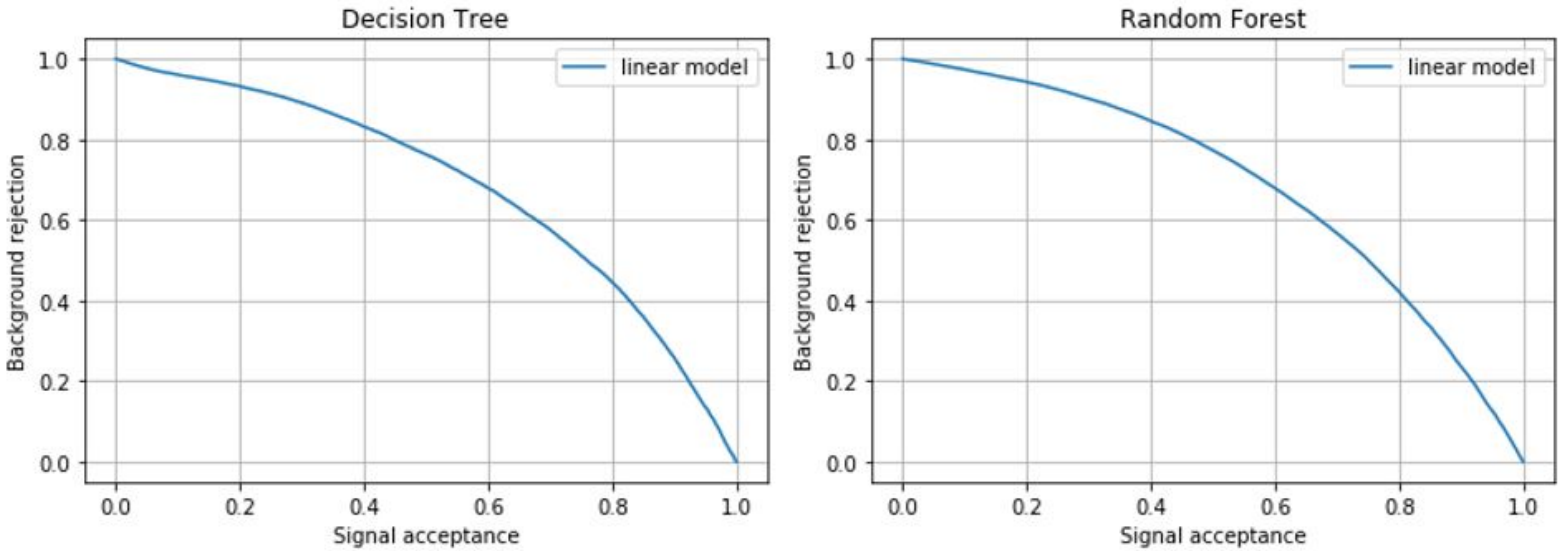
Сравнение распределений некоторых параметров TimeAtHeightX для синтезированных (Minimax GAN) и калибровочных фоновых событий для детектора ANG5

Minimax GAN, входные данные - импульсы заряда

Результаты, полученные для детектора ANG5 при обучении классификаторов на синтетических данных (Minimax GAN) и их проверке на калибровочных данных

	accuracy	roc_auc	f1
Decision Tree	0.64 ± 0.02	0.64 ± 0.02	0.64 ± 0.02
Random Forest	0.64 ± 0.02	0.64 ± 0.02	0.64 ± 0.02

Minimax GAN, входные данные - импульсы заряда



ROC-кривые для детектора ANG5 при обучении классификаторов на синтетических данных (Minimax GAN) и их проверке на калибровочных данных. Background rejection - TNR (1-FPR), signal efficiency - TPR

Minimax GAN, входные данные - импульсы заряда

Таблица 4 – Результаты, полученные для детектора ANG5 при разделении синтетических (Minimax GAN) и реальных данных

	accuracy	roc_auc	f1
Decision Tree	0.66 ± 0.02	0.66 ± 0.02	0.65 ± 0.02
Random Forest	0.76 ± 0.02	0.76 ± 0.02	0.76 ± 0.02

Minimax GAN, входные данные - импульсы заряда

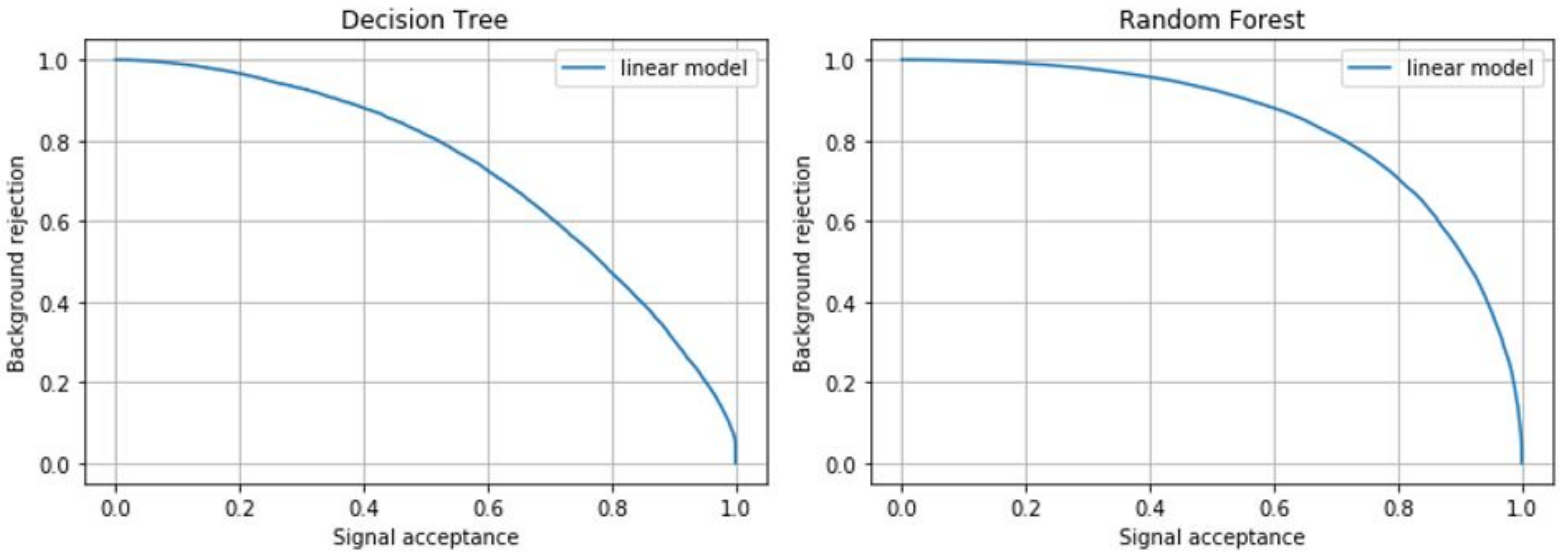


Рисунок 17 – ROC-кривые для детектора ANG5 при разделении синтетических (Wasserstein GAN) и реальных данных. Background rejection - TNR (1-FPR), signal efficiency - TPR