

Выпускная квалификационная работа магистра

**Моделирование сигналов полупроводниковых
германиевых детекторов с помощью нейронных
сетей для эксперимента LEGEND**

Студент: Зарецкий Н. Д.

Научный руководитель: д.ф.-м.н., проф. Скорохватов М. Д.

Научный консультант: Черногоров А. Е.

Москва, 2023

Введение

- **Мотивация:** моделирование с помощью нейронных сетей может стать дополнением к Монте-Карло моделированию или его альтернативой.
- **Цель работы:** моделирование сигнала для ICPC детекторов в эксперименте LEGEND с помощью нейронных сетей.

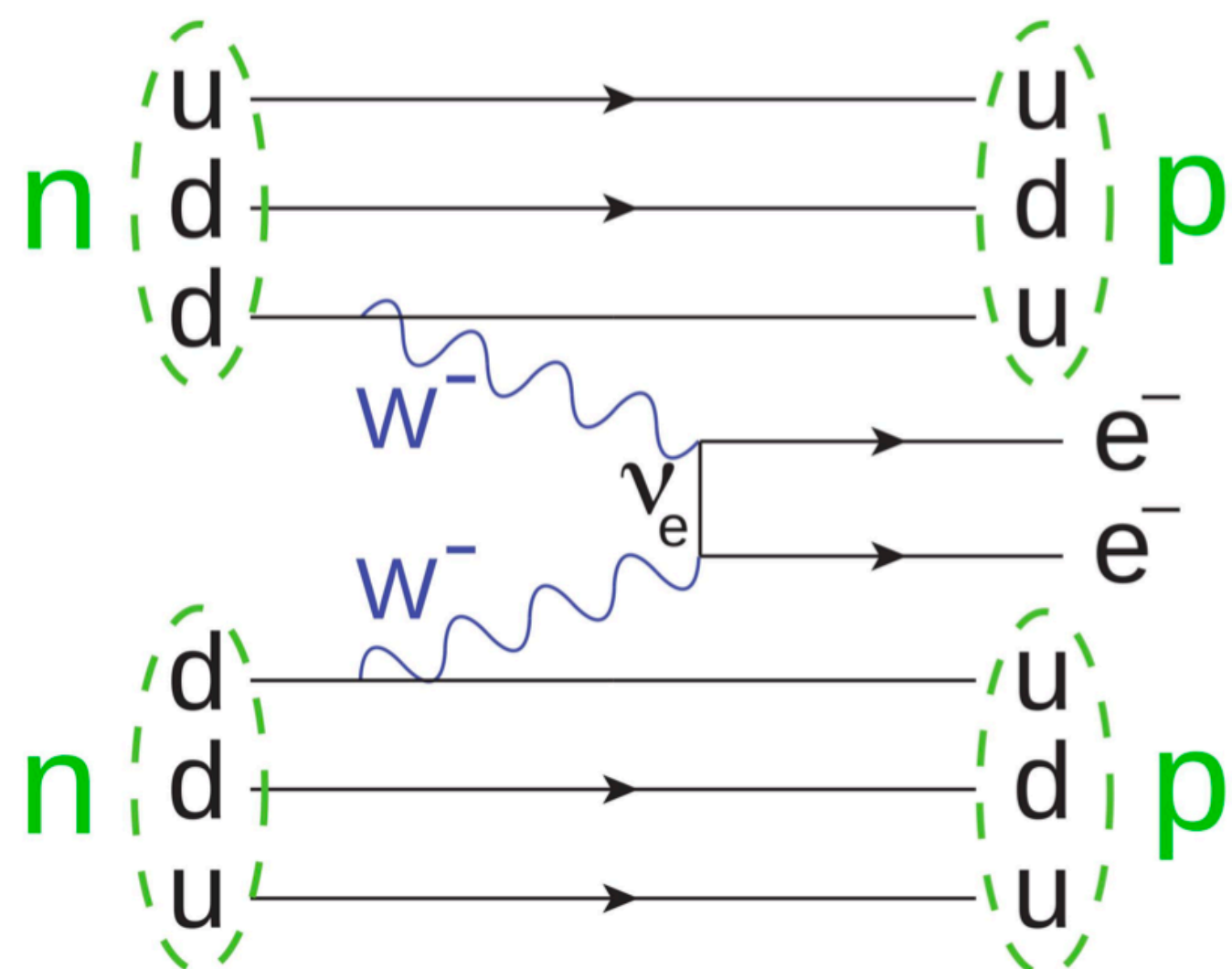


Рисунок 1 - Диаграмма $0\nu\beta\beta$

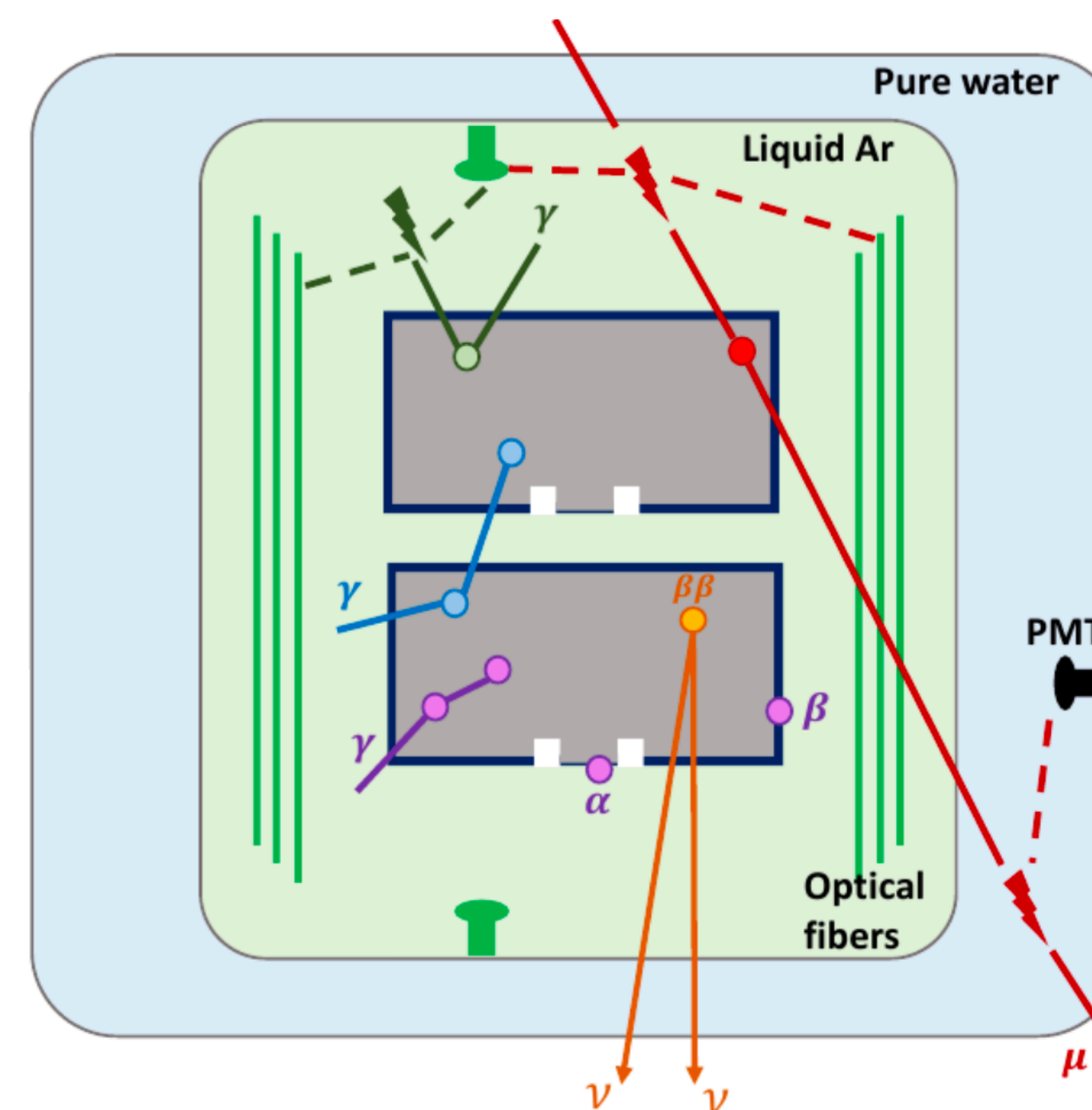


Рисунок 2 - Схема эксперимента LEGEND

Калибровка детекторов

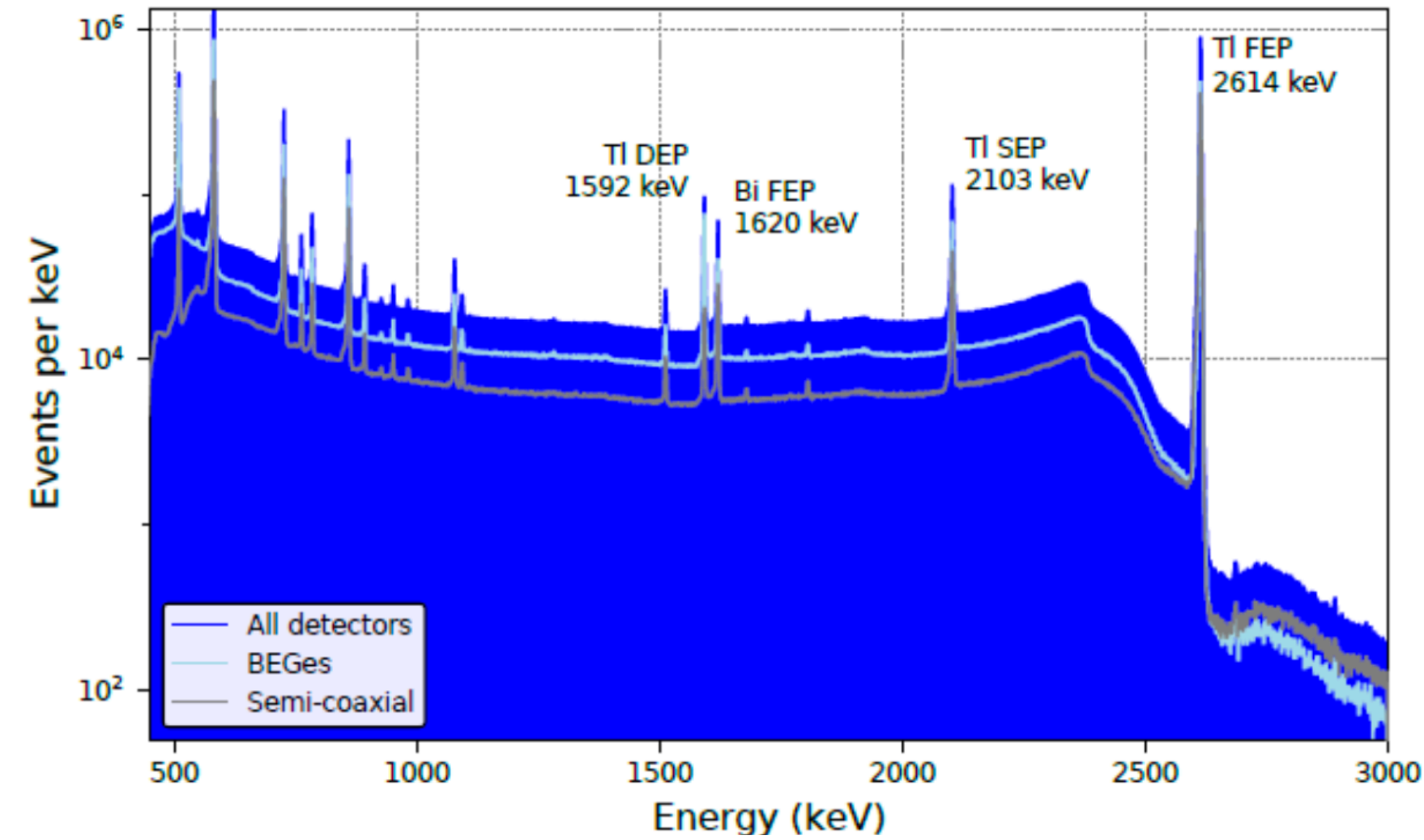


Рисунок 3 - Энергетический спектр ^{228}Th

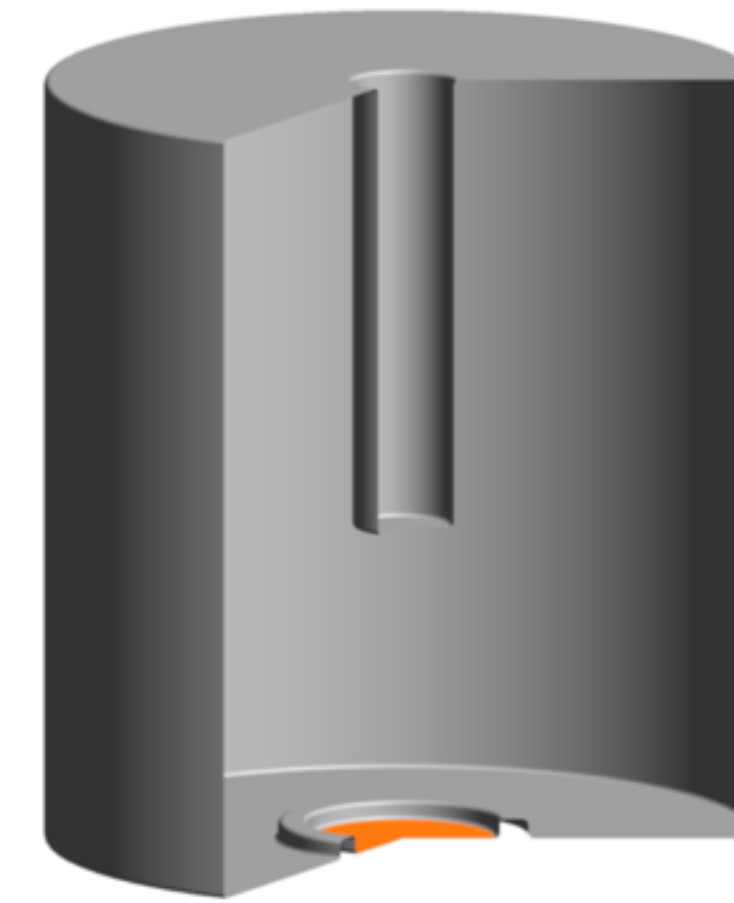


Рисунок 4 - Схема ICPC детектора

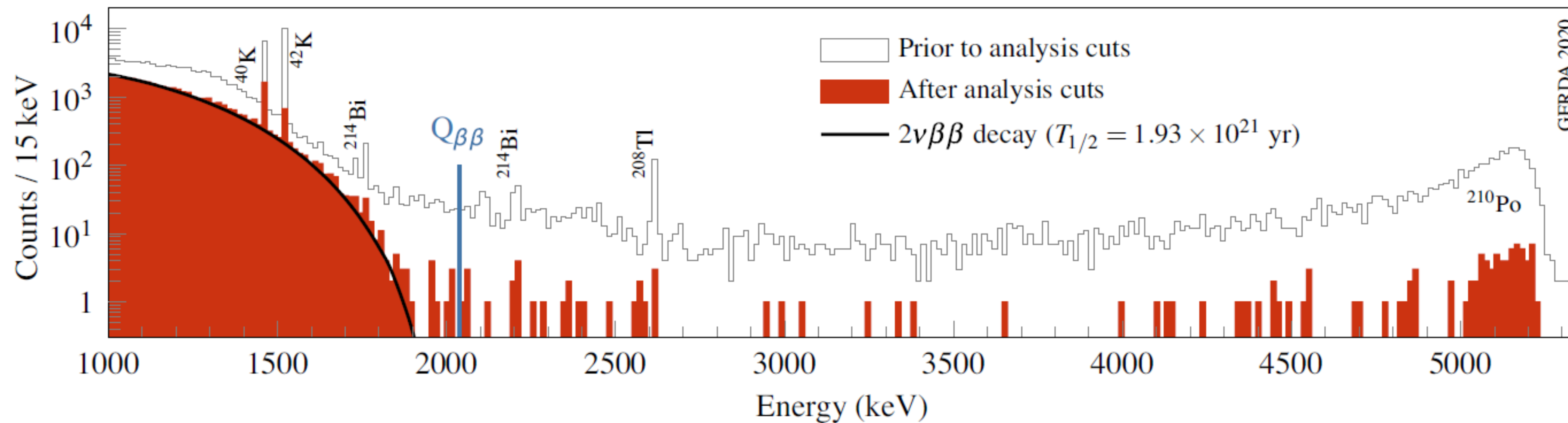


Рисунок 5 - Энергетический спектр эксперимента GERDA

Генеративно-сопоставительная сеть (GAN)

- GAN применяются для генерации новых данных, которые могли бы правдоподобно присутствовать в оригинальном наборе
- GAN состоят из двух подмоделей (генератора и дискриминатора), обучающихся одновременно
- Задача оптимизации: $\min_G \max_D \{ \log(D(x)) + \log(1 - D(G(z))) \}$, где z - шум, поданный на вход генератору; $D(x)$ - «вероятность», выданная дискриминатором в процессе классификации; $G(z)$ - данные, синтезированные генератором

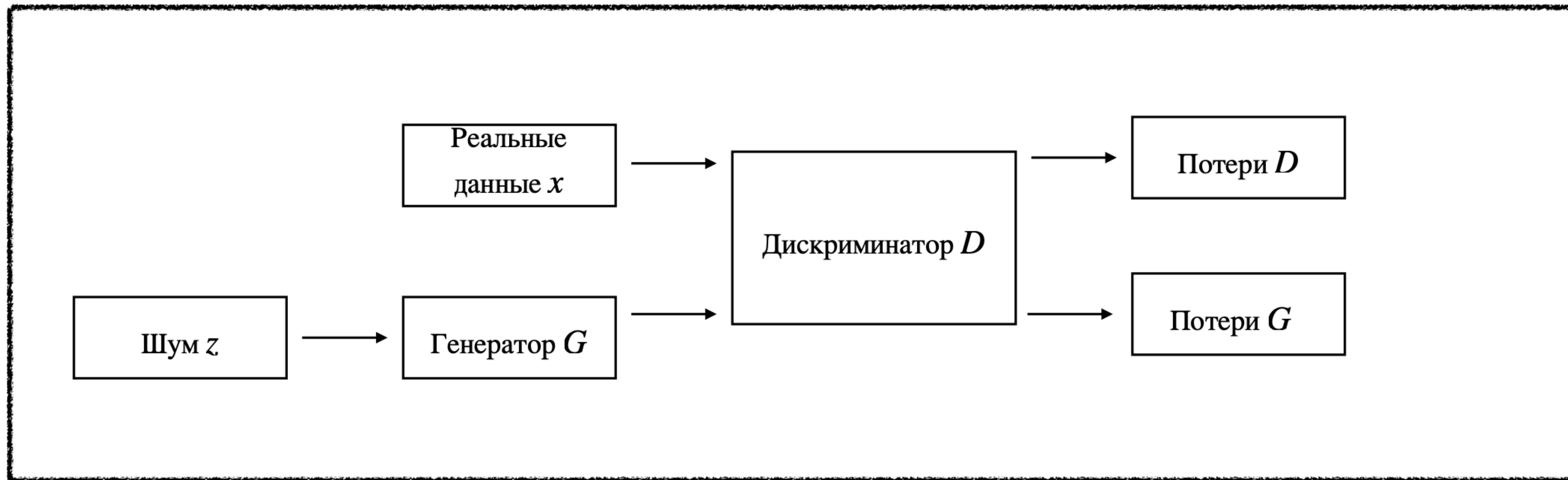
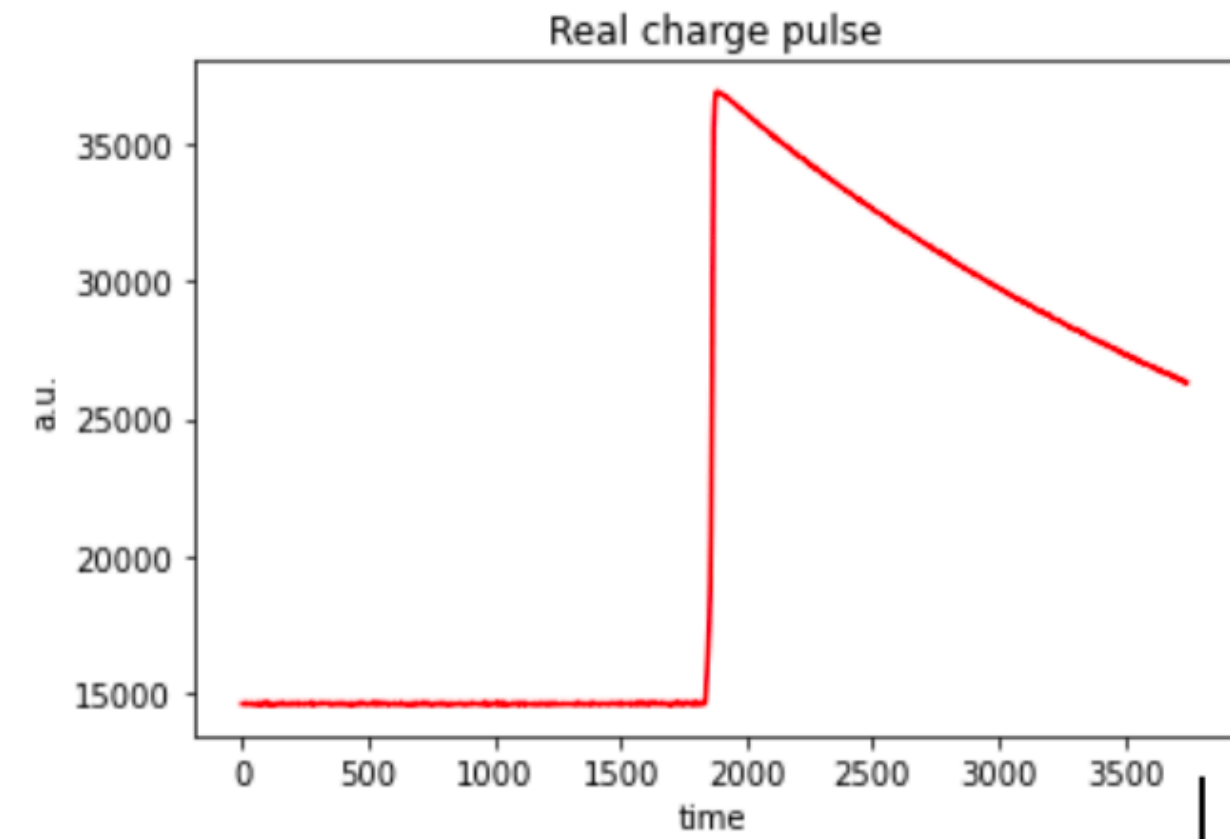
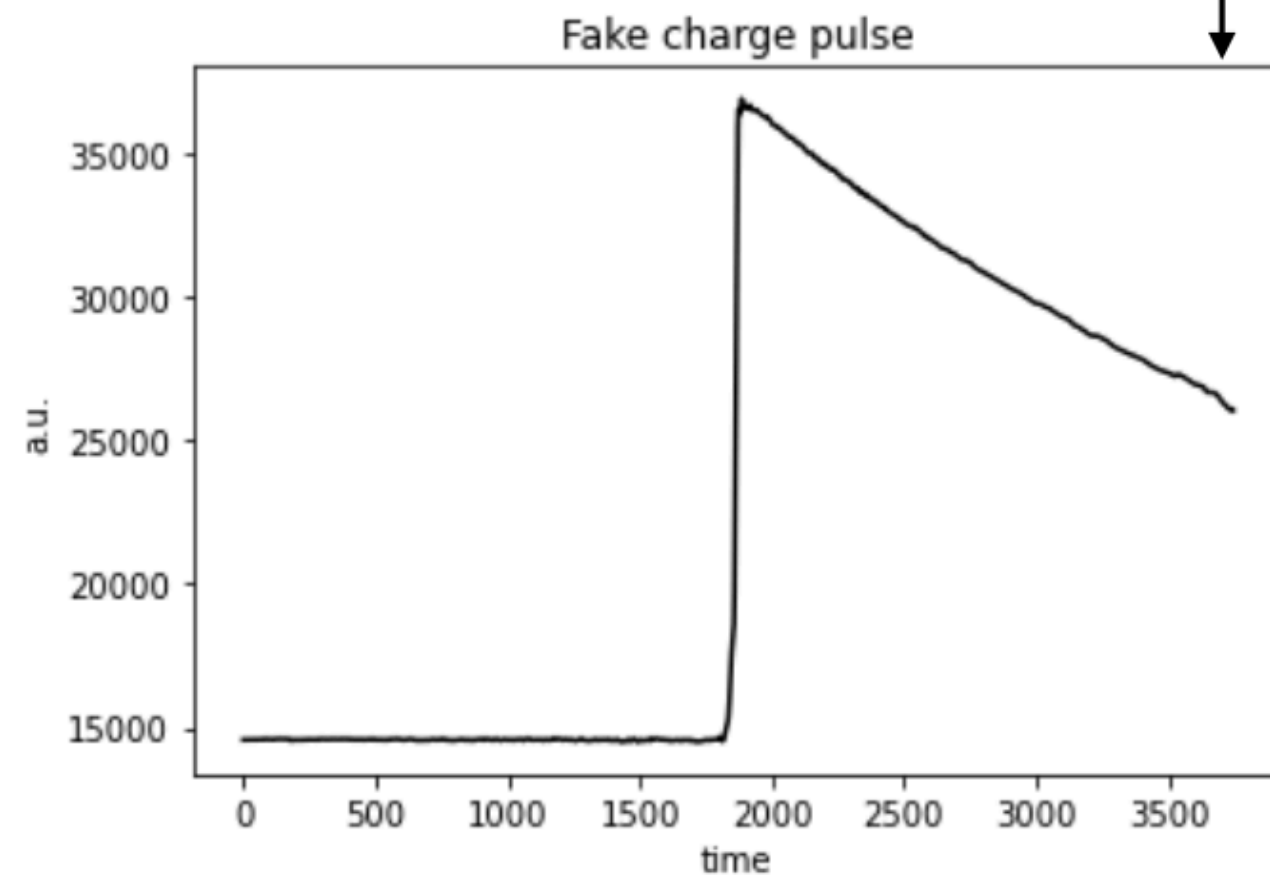


Рисунок 6 - Схема работы генеративно-сопоставительной сети

Входные данные



Входные данные



Результат

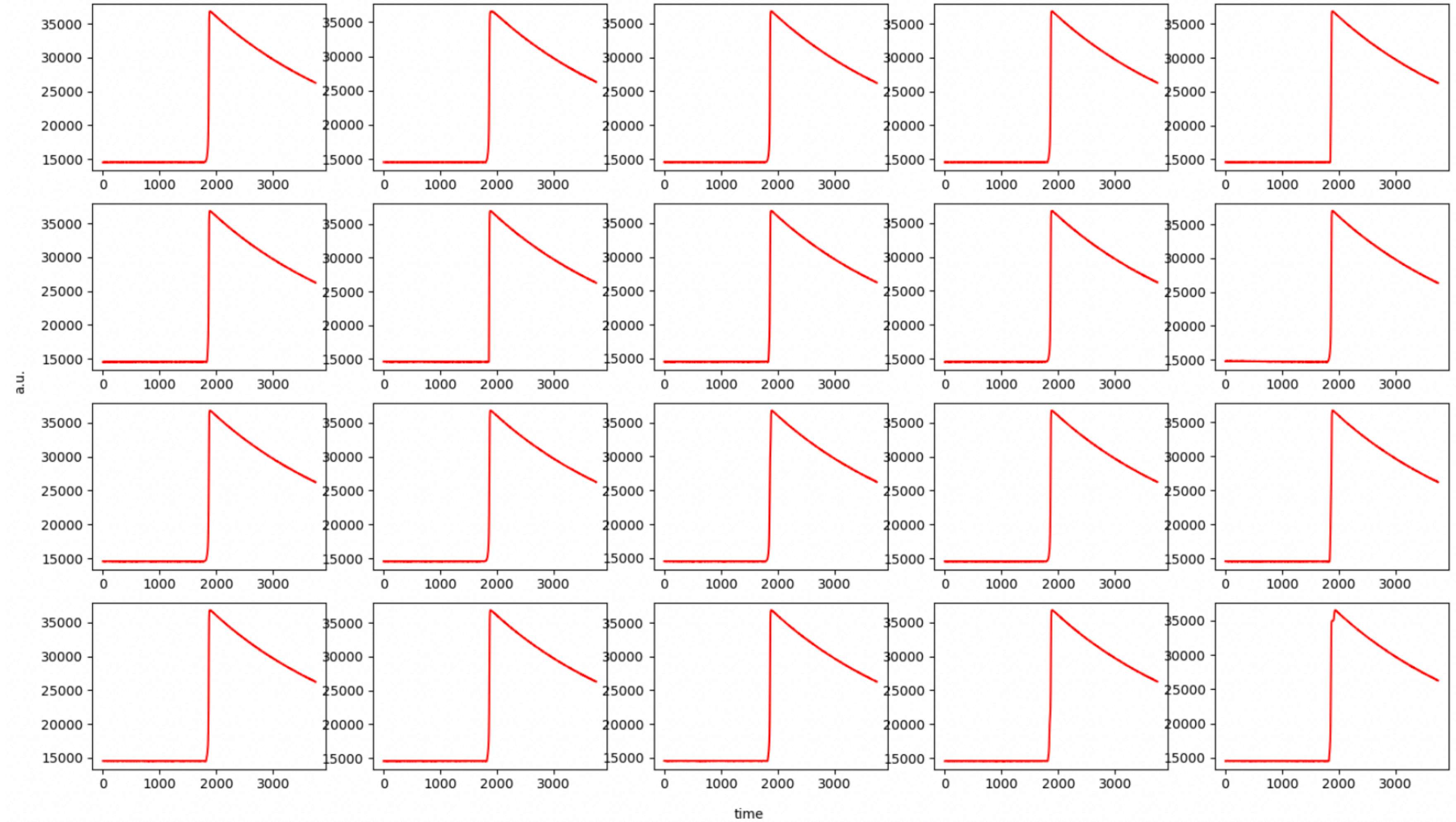
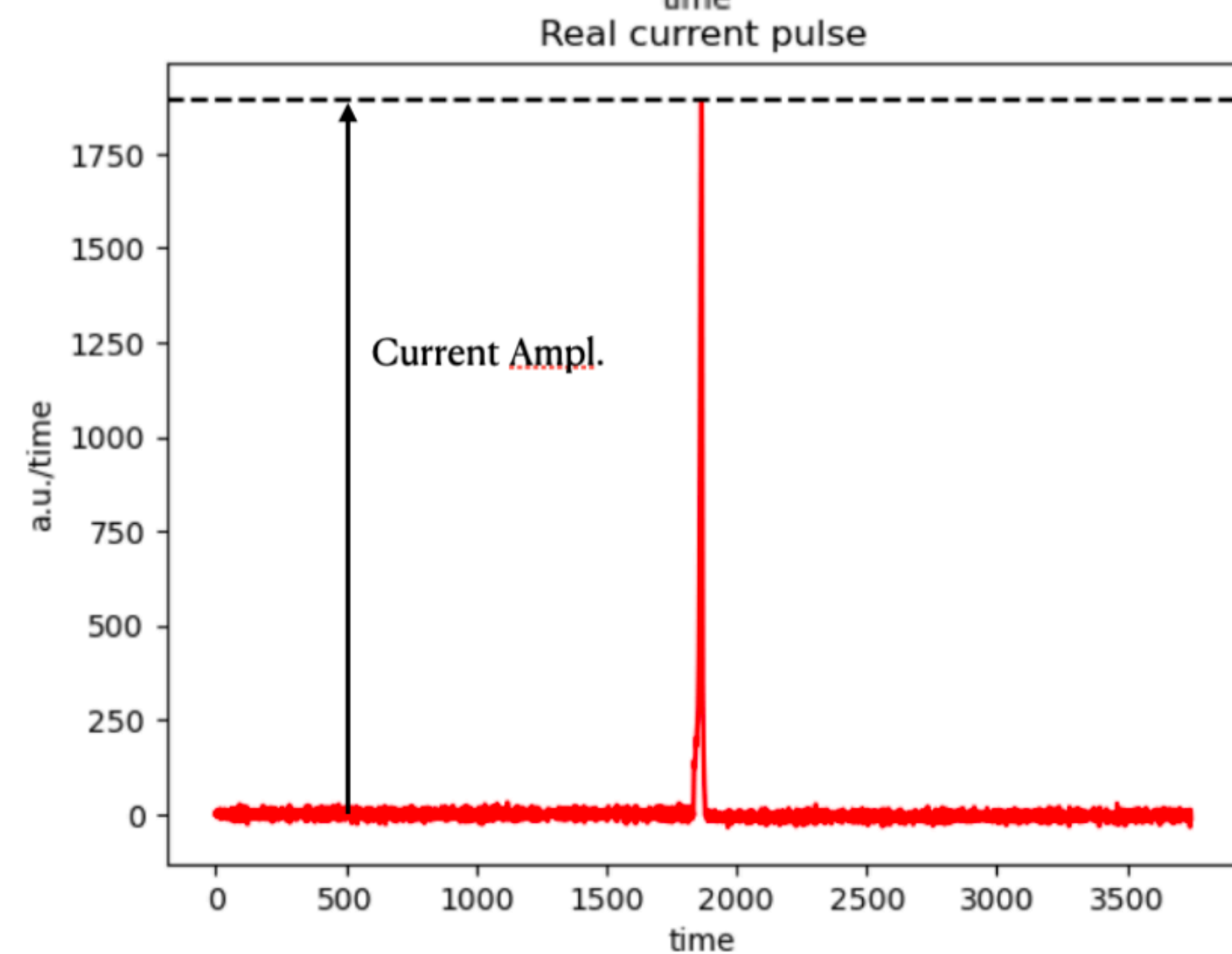
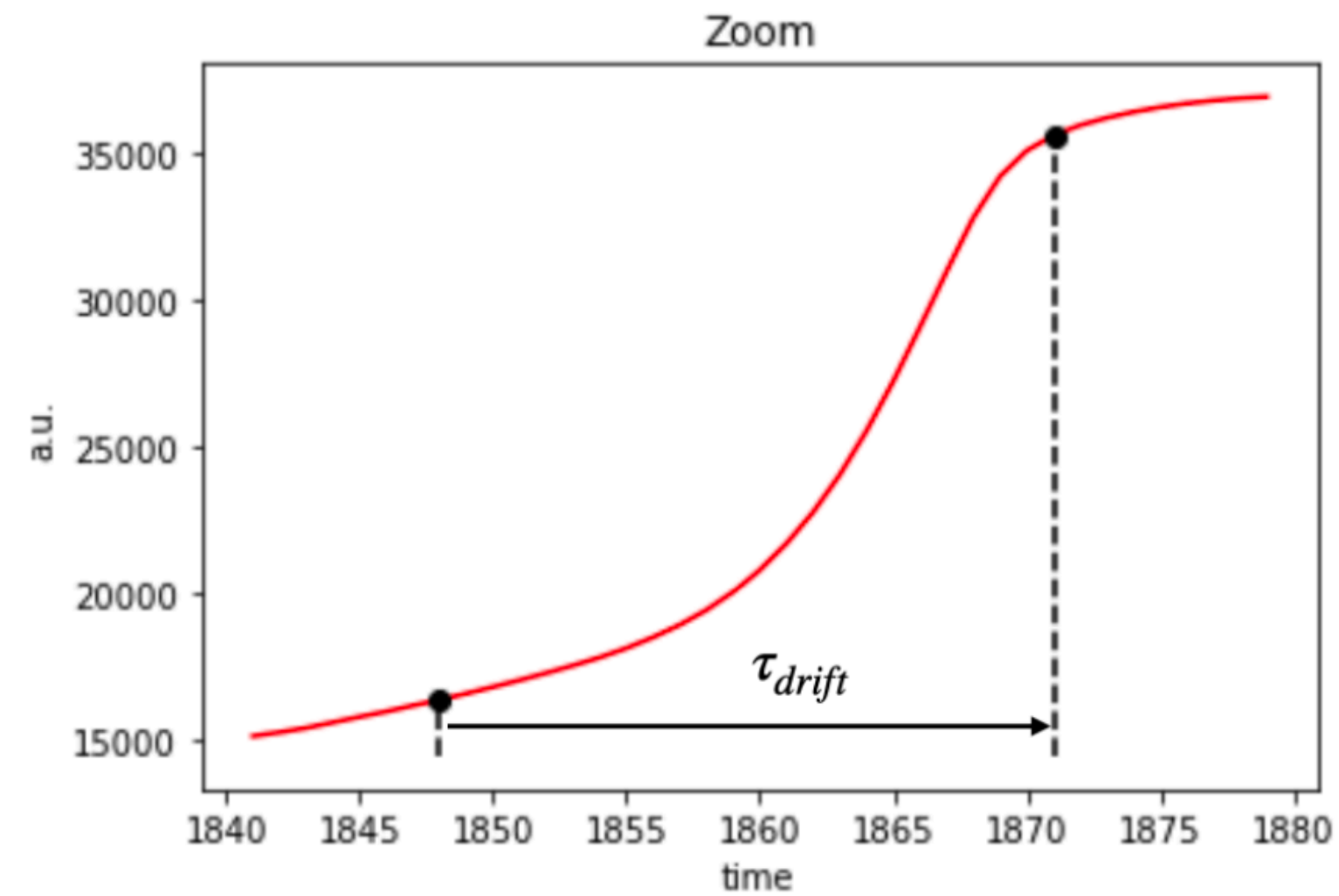
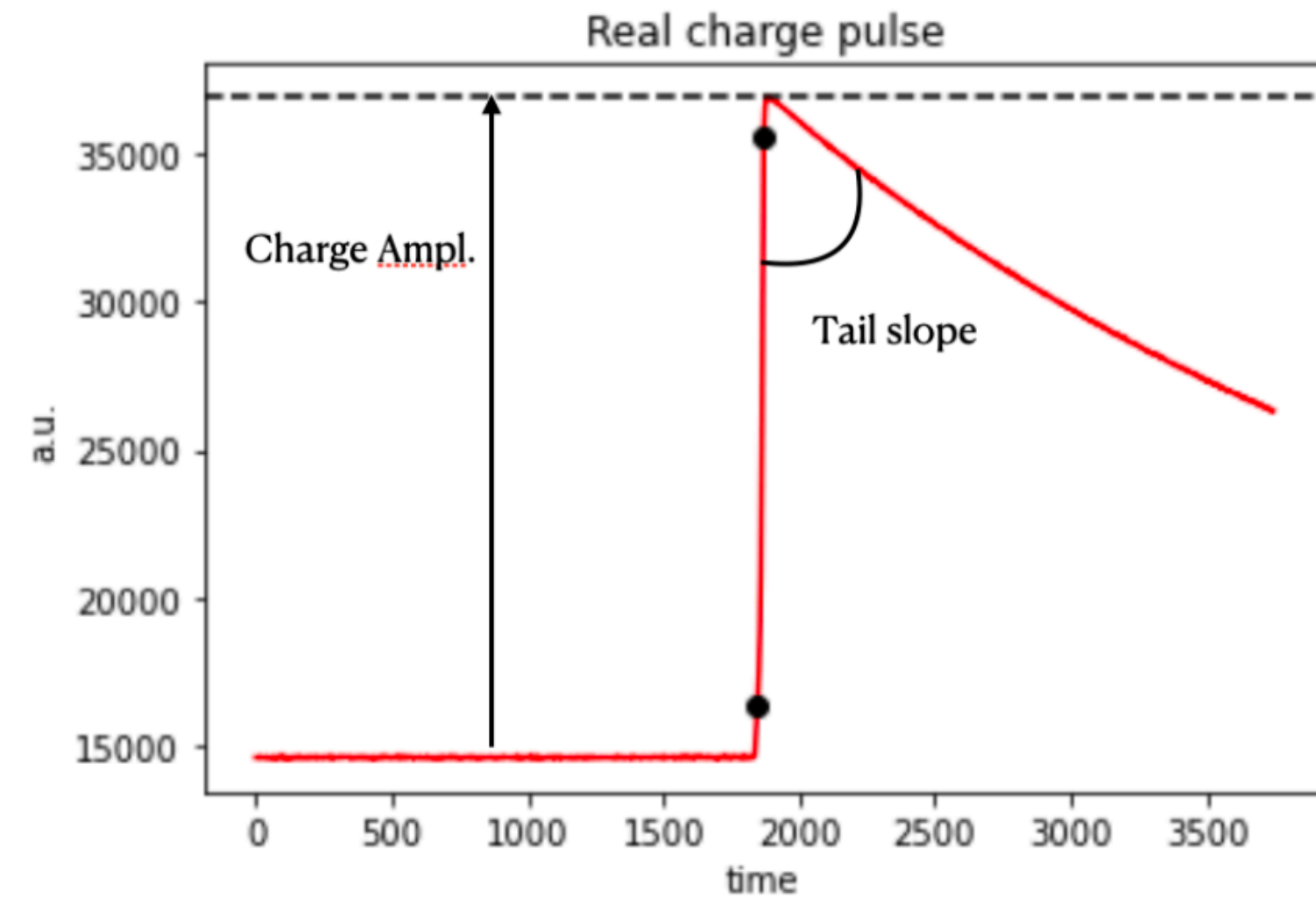


Рисунок 7 - Схема работы генеративной модели

Рисунок 8 - Пример входных данных

Оценка качества смоделированных сигналов



Для сравнения синтезированных данных с реальными были использованы 4 параметра:

- Амплитуда зарядового импульса
- Амплитуда токового импульса
- Угол наклона хвостовой части зарядового импульса
- Время дрейфа τ_{drift}

Рисунок 9 - Параметры, используемые для оценки правдоподобия сгенерированных сигналов

GAN

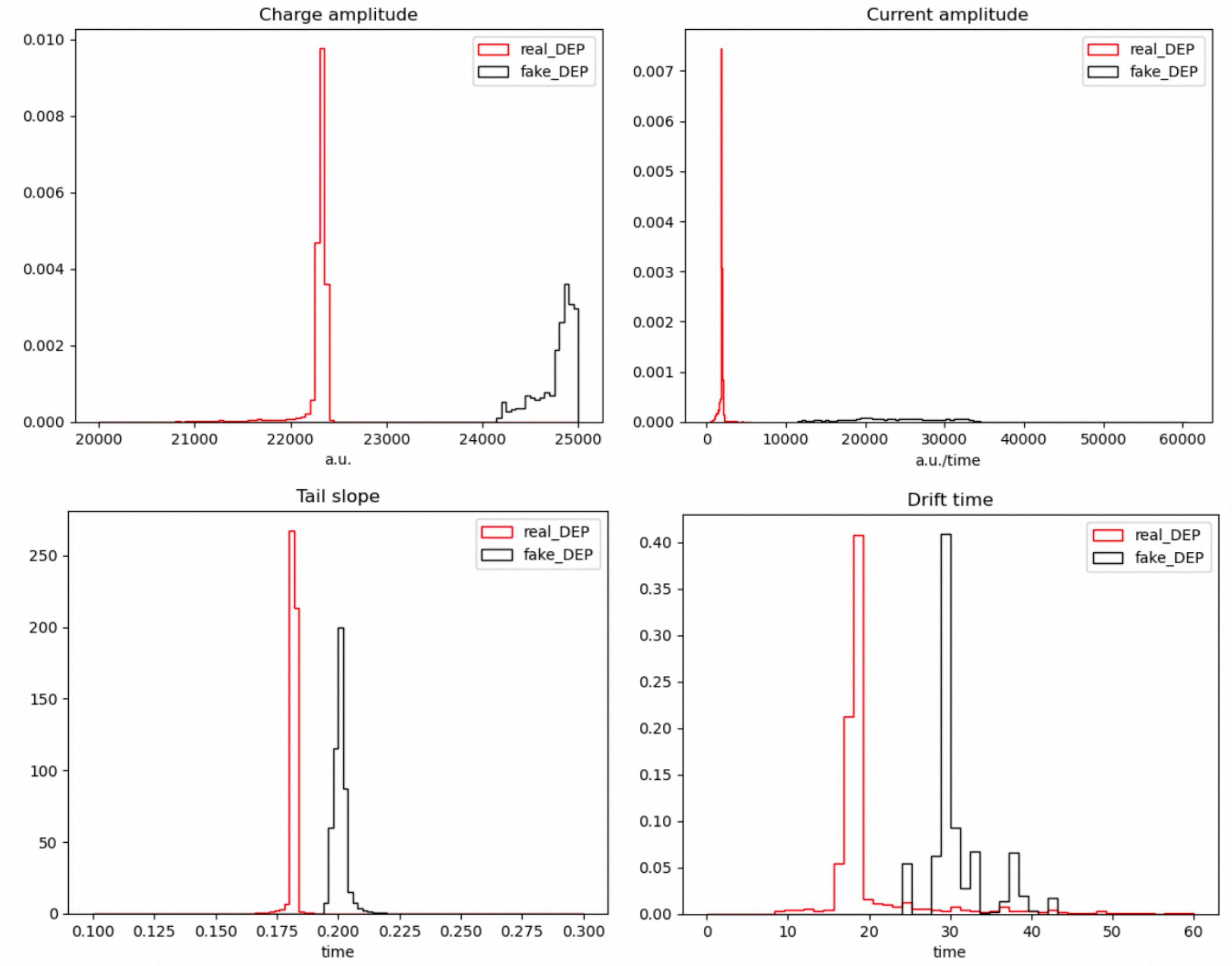
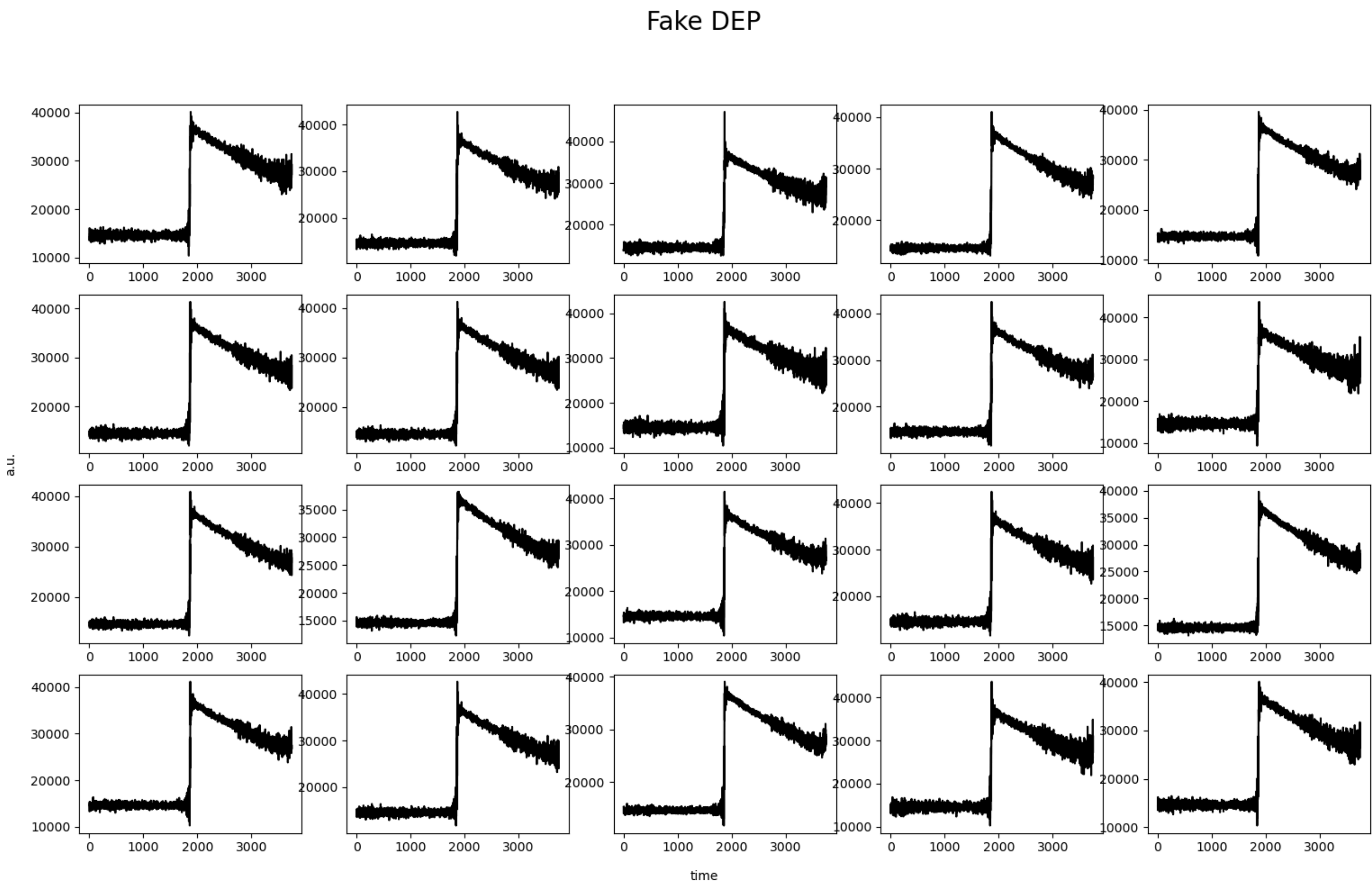


Рисунок 10 - Примеры импульсов, полученных в результате обучения модели GAN

Рисунок 11 - Распределения параметров импульсов, полученных в результате обучения модели GAN

GAN с использованием метода главных компонент (PCA)

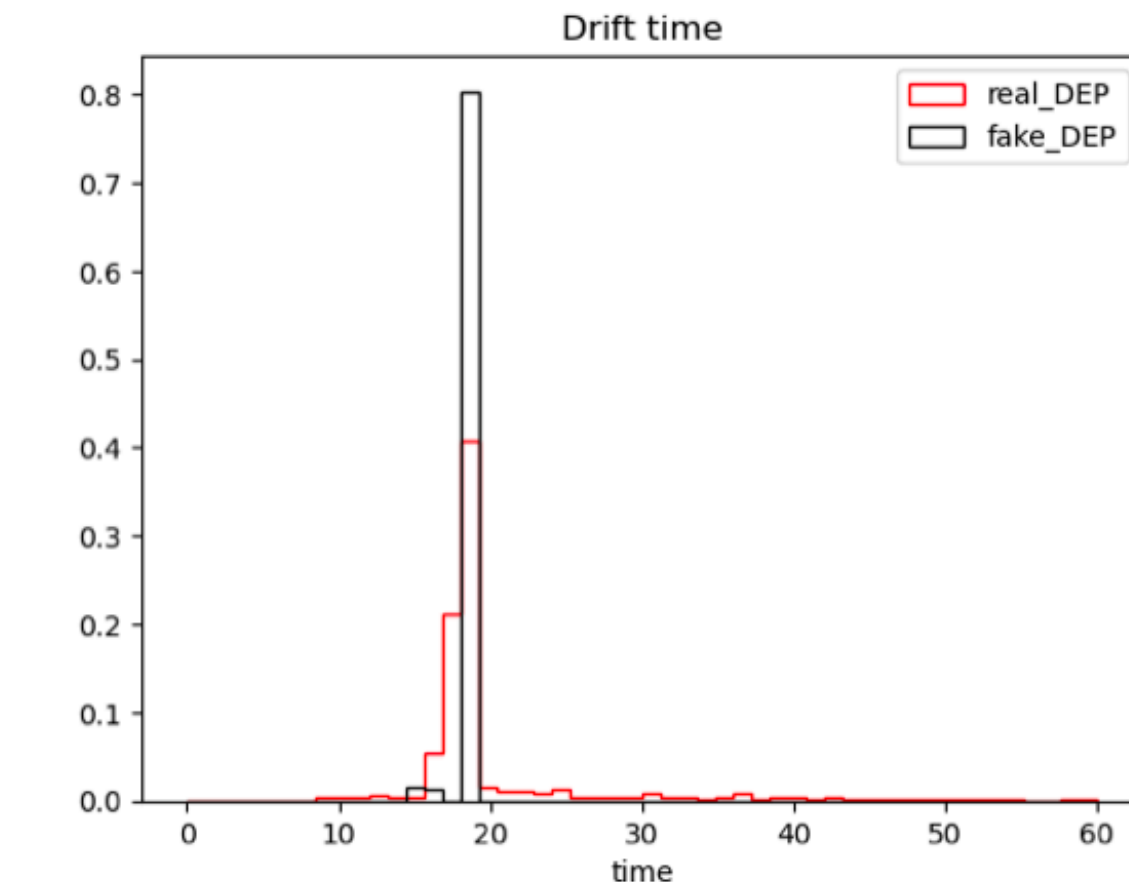
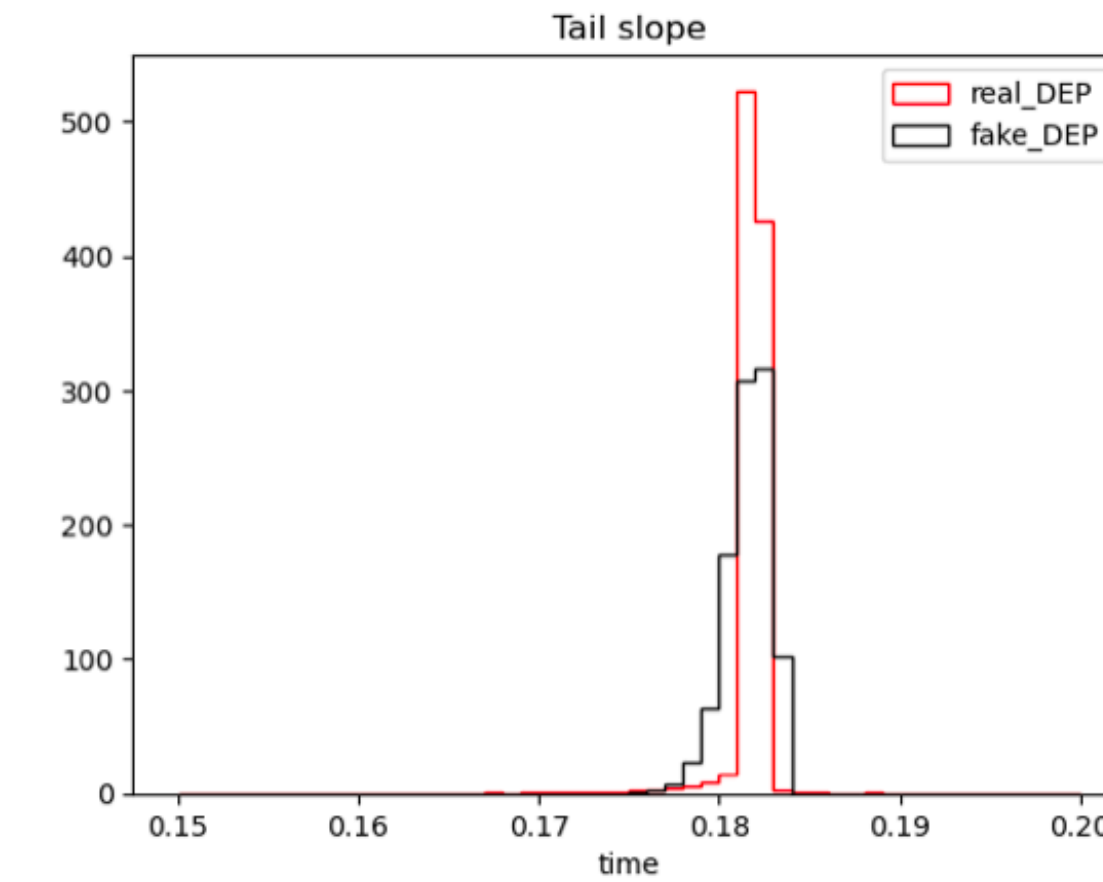
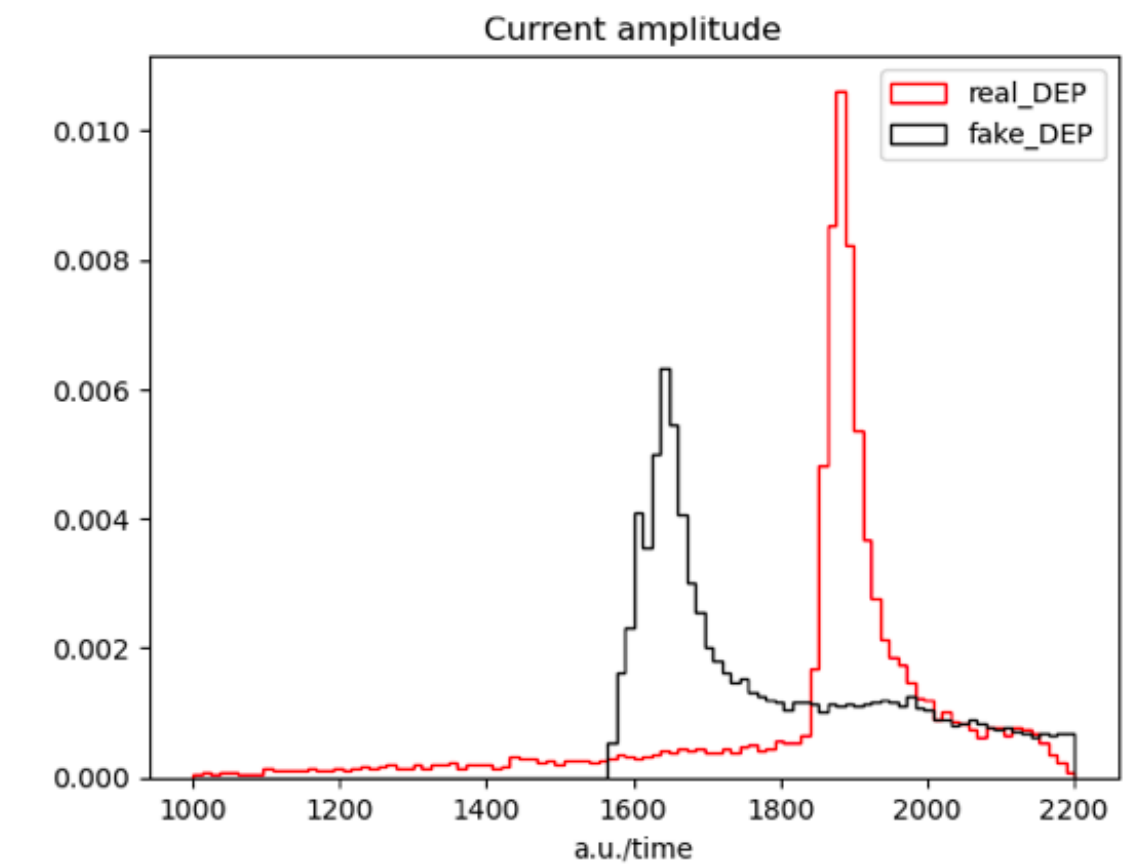
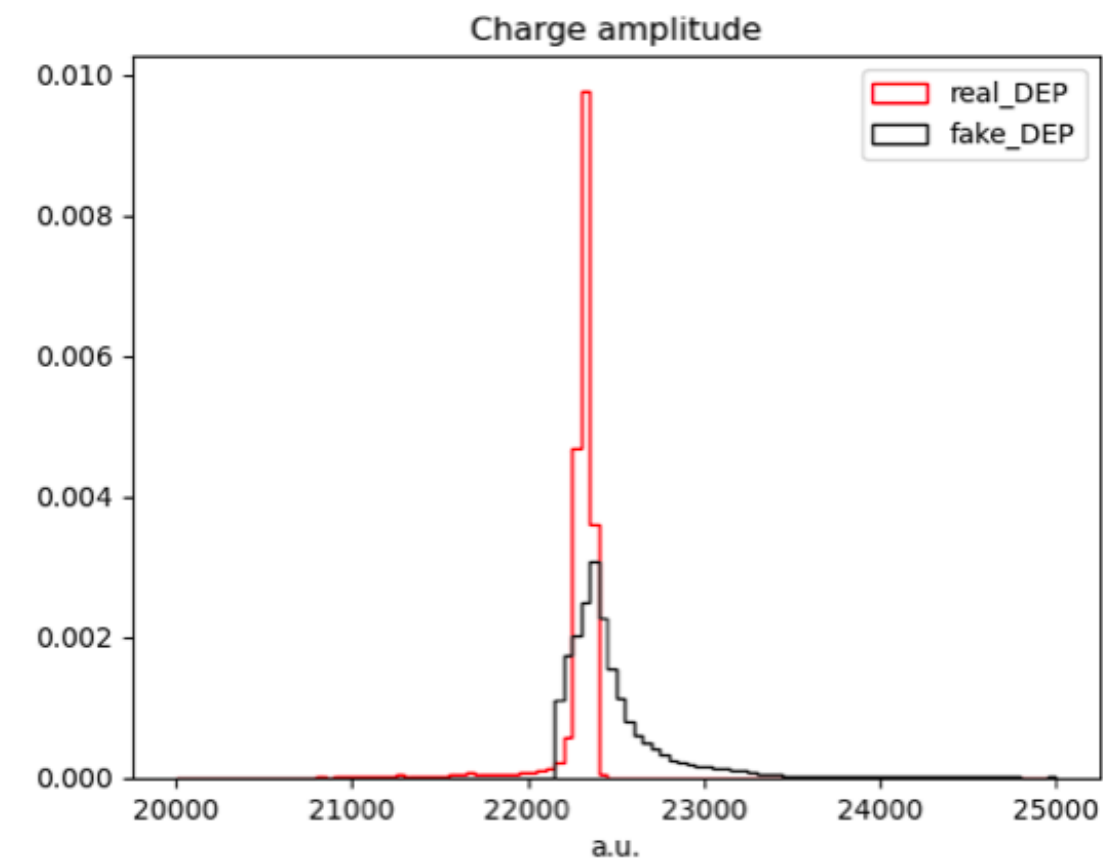
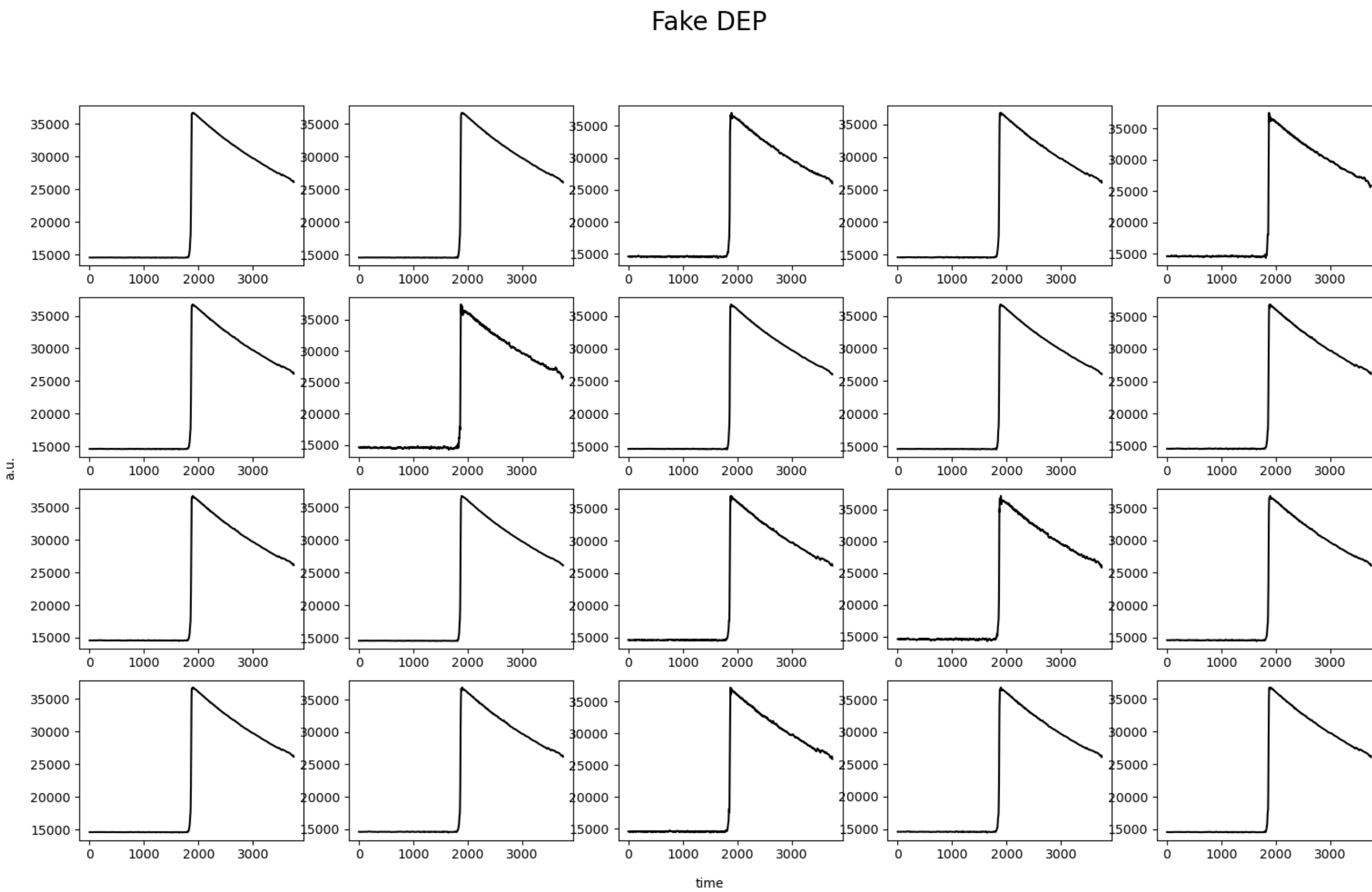
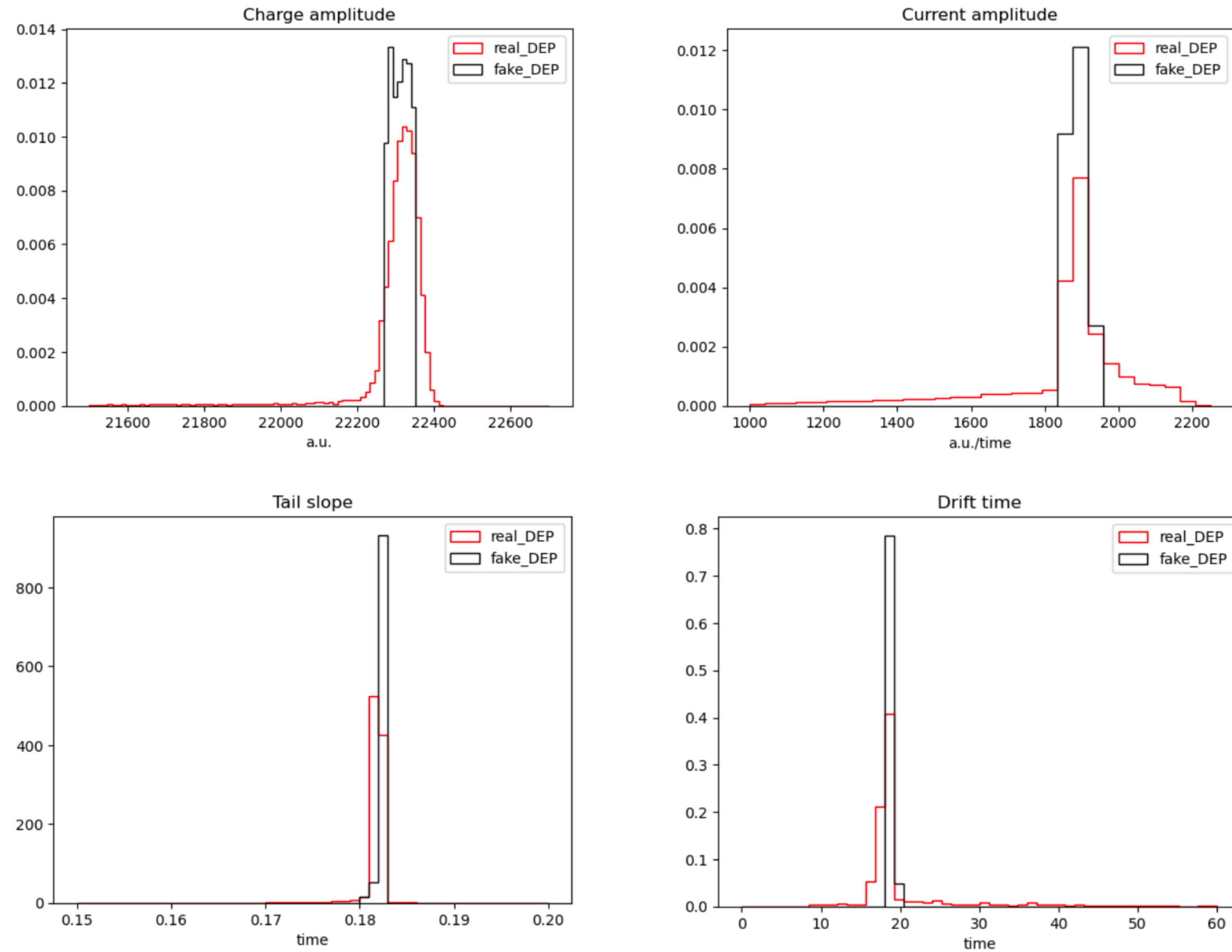


Рисунок 12 - Примеры импульсов, полученных в результате обучения GAN с PCA

Рисунок 13 - Распределения параметров импульсов, полученных в результате обучения GAN с PCA

Критерии отбора



Критерии отбора:

- $22270 \text{ a.u.} < \text{Charge amplitude} < 22350 \text{ a.u.}$
- $1850 \text{ a.u./time} < \text{Current amplitude} < 2000 \text{ a.u./time}$

Рисунок 14 - Распределения параметров импульсов,
полученных в результате обучения GAN с PCA
после применения критериев отбора

Обобщение. Проверка модели GAN-PCA на данных, полученных на другом детекторе

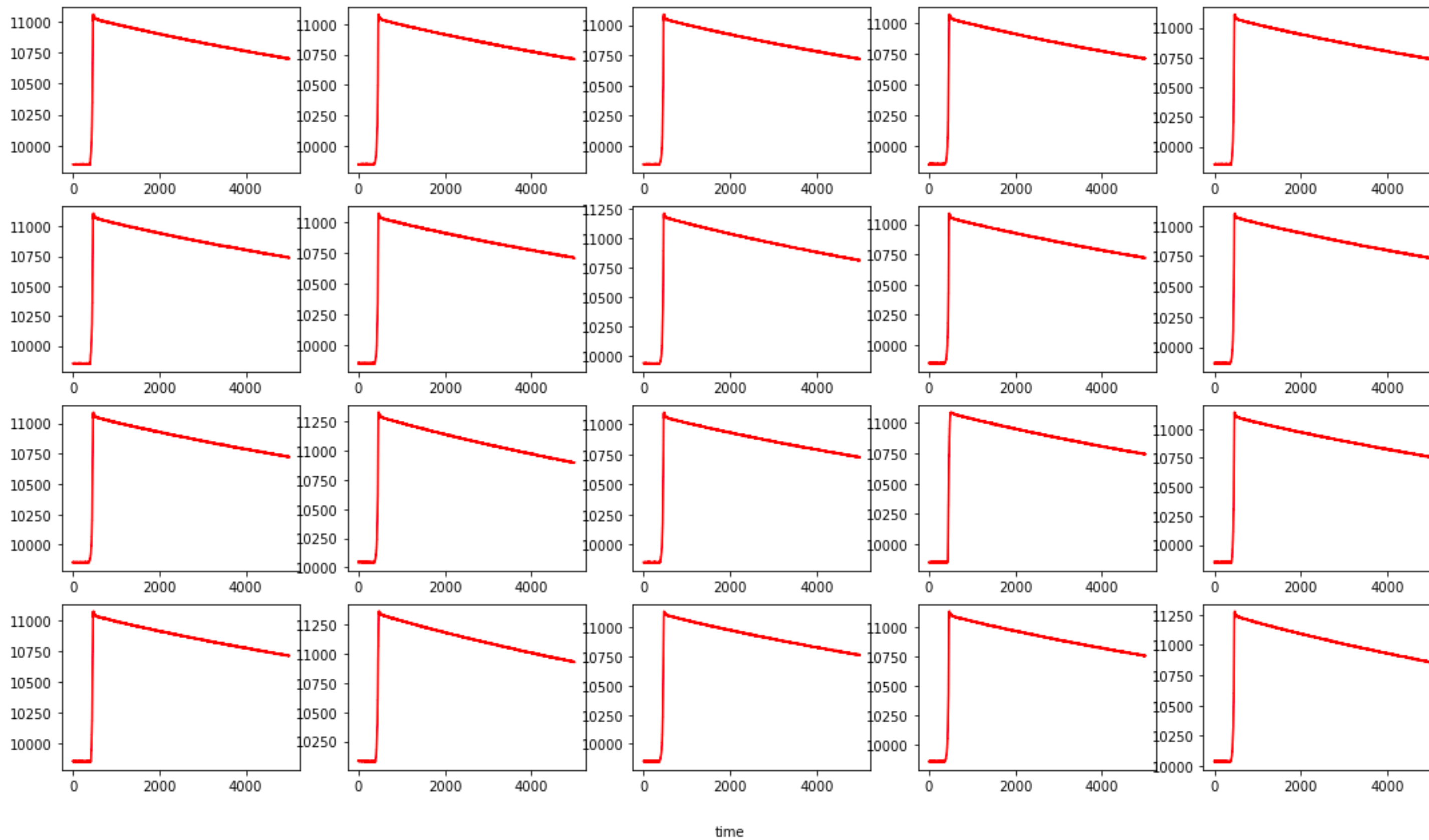


Рисунок 15 - Пример входных данных

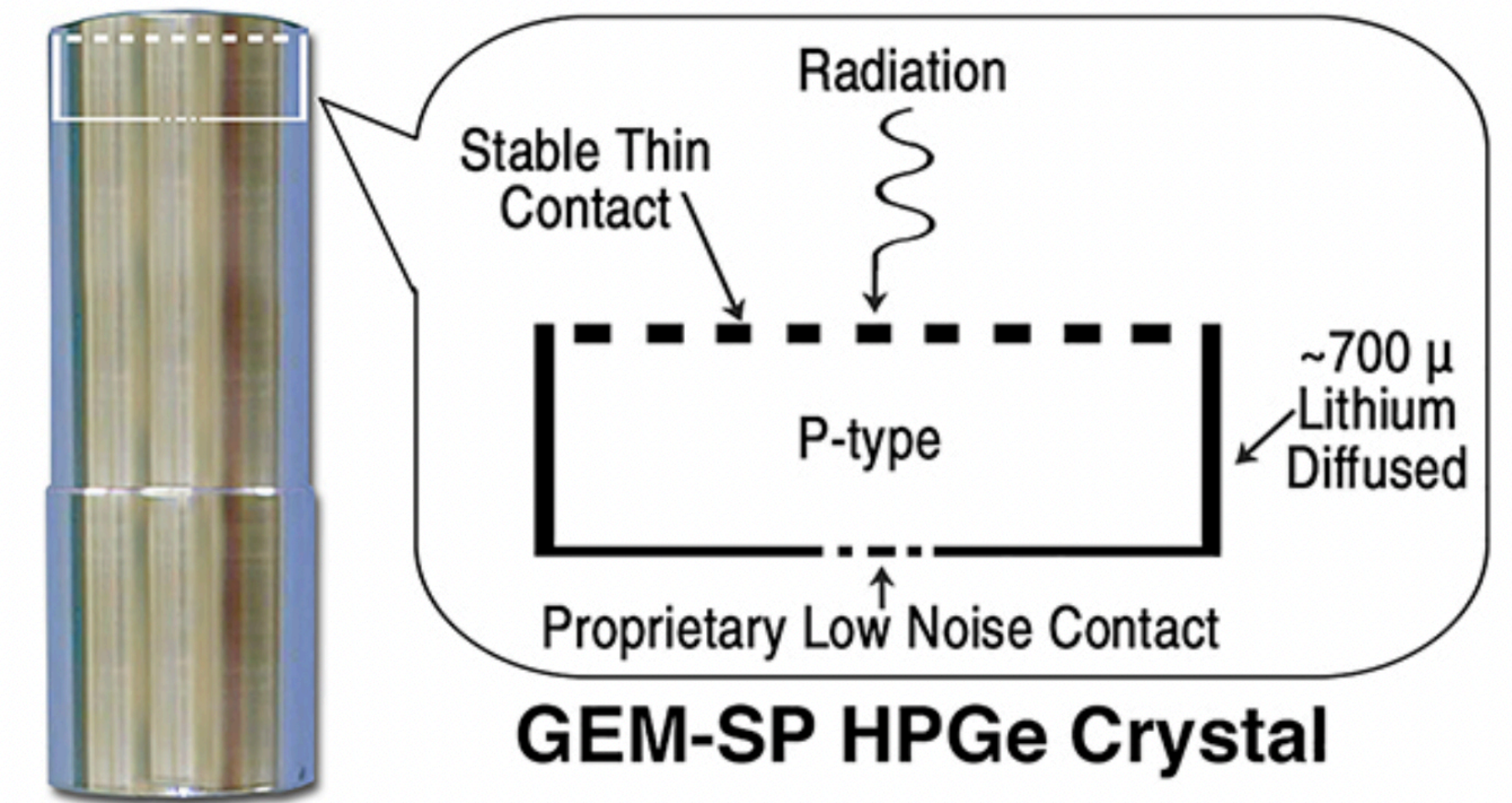


Рисунок 16 - Детектор SP5020 ORTEC

GAN с использованием PCA

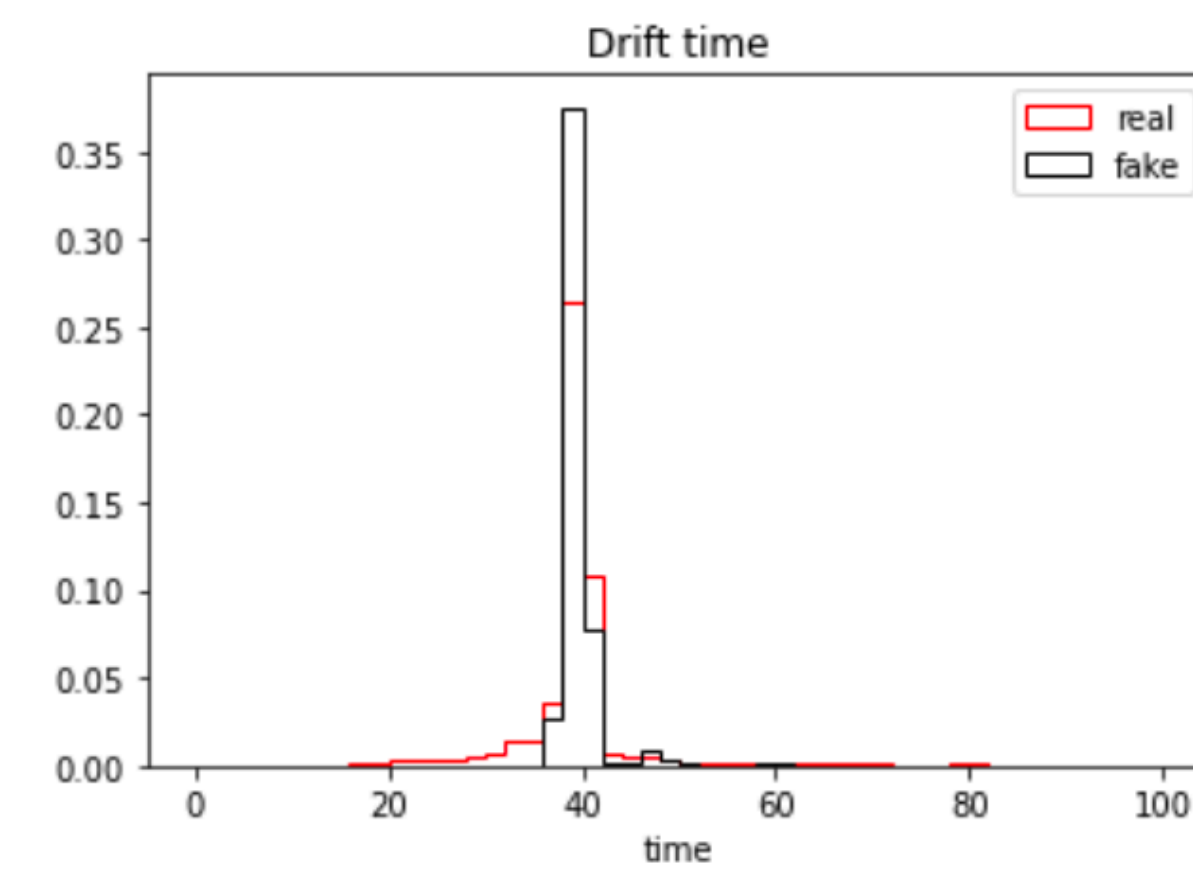
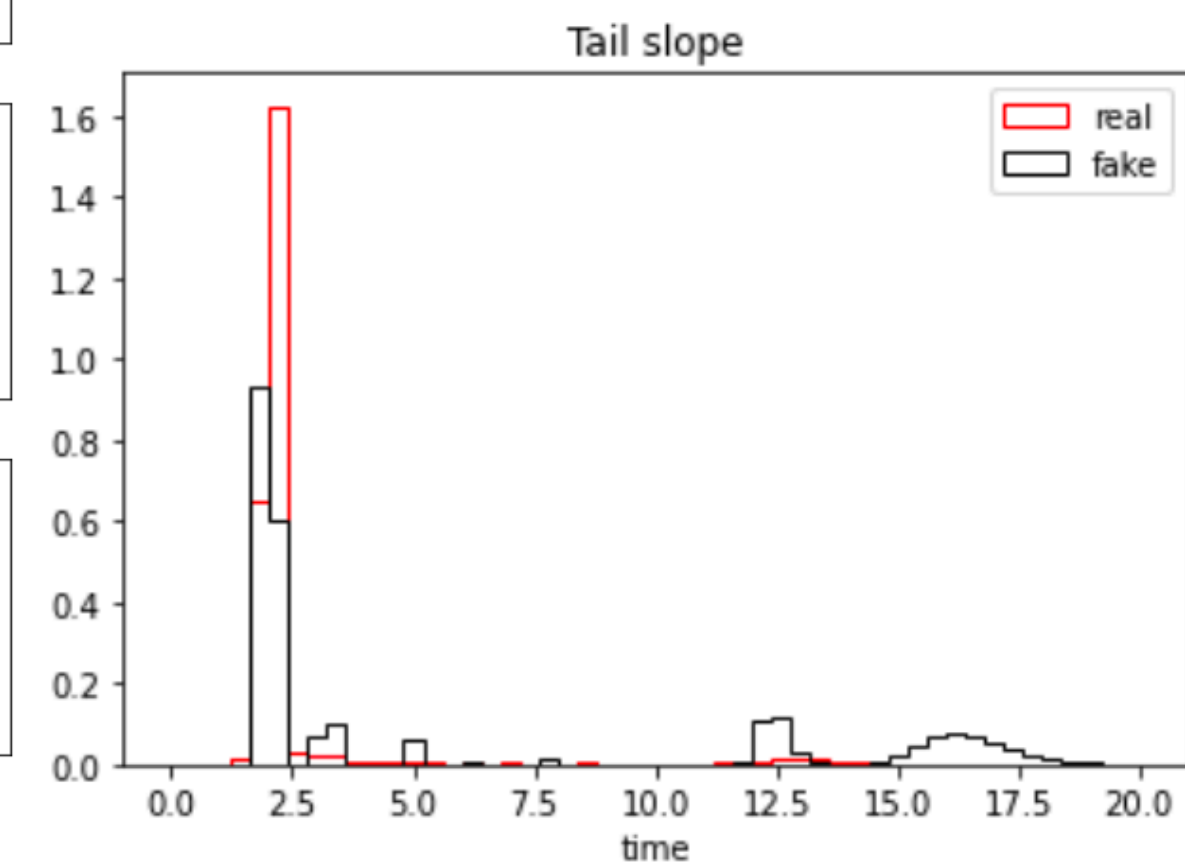
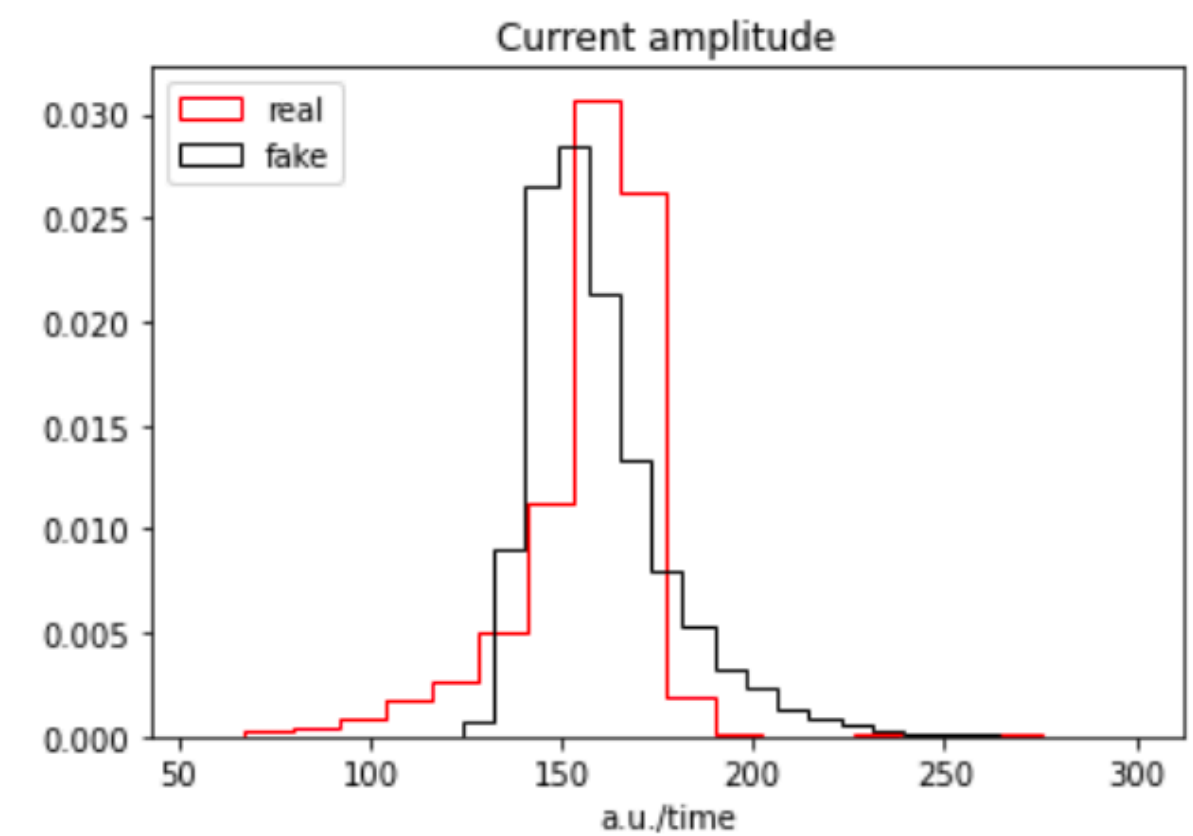
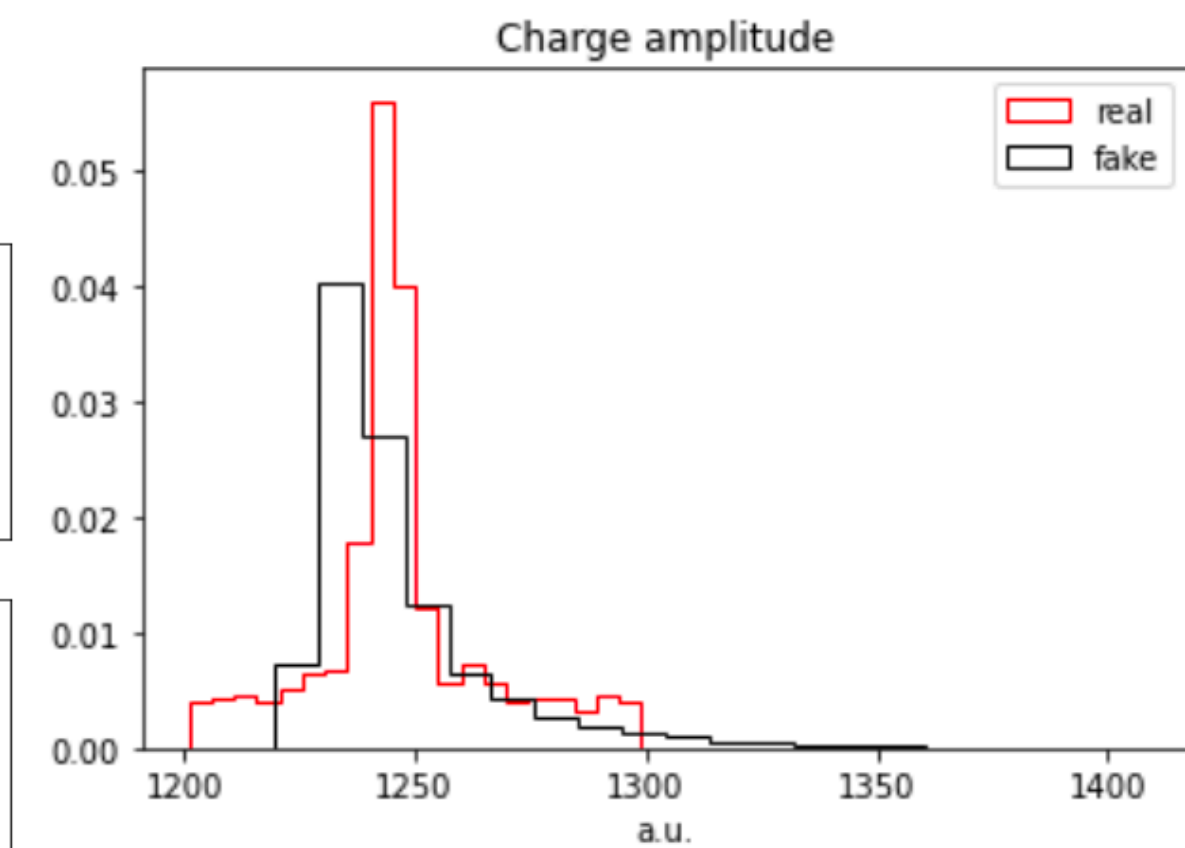
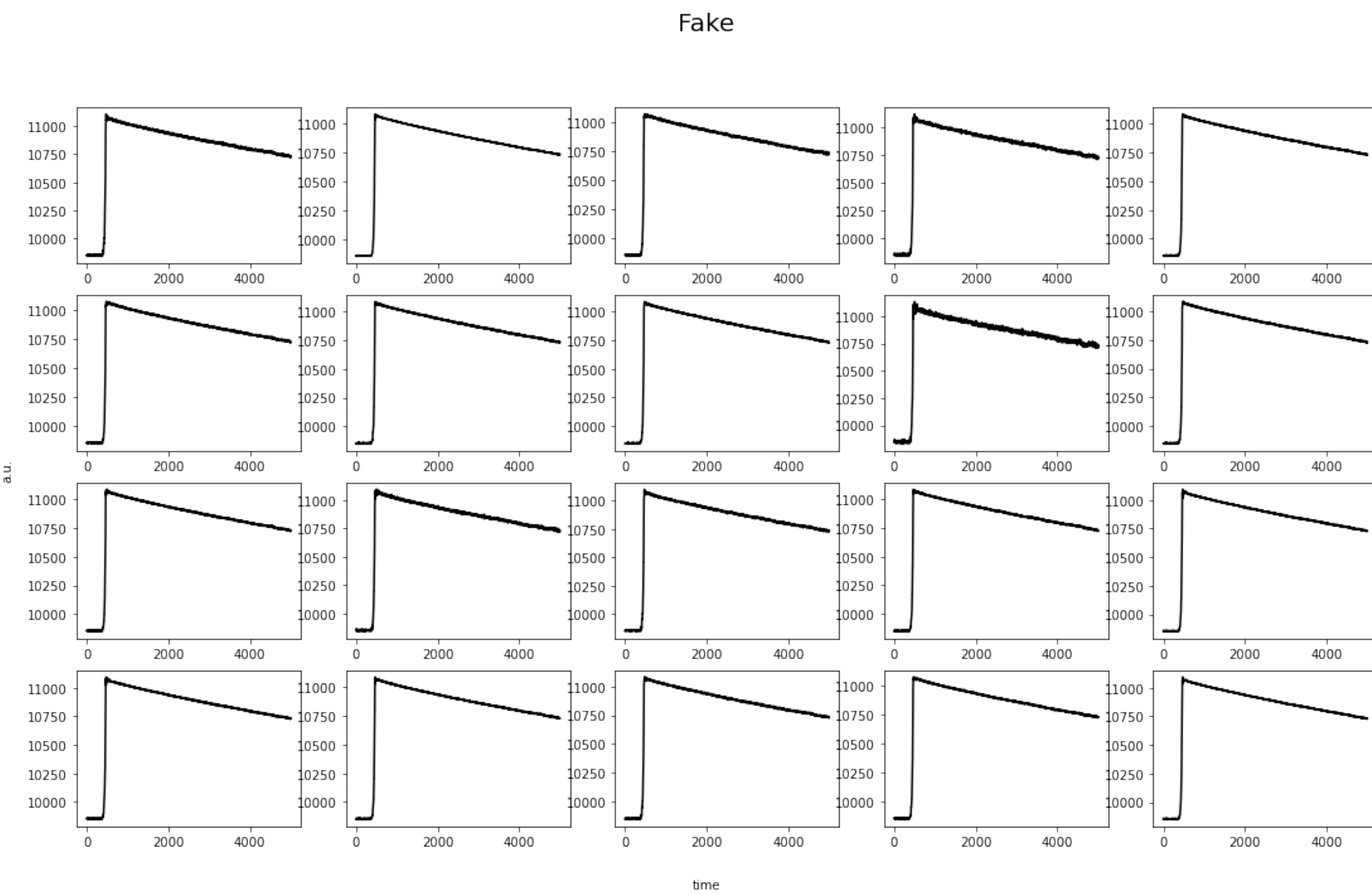
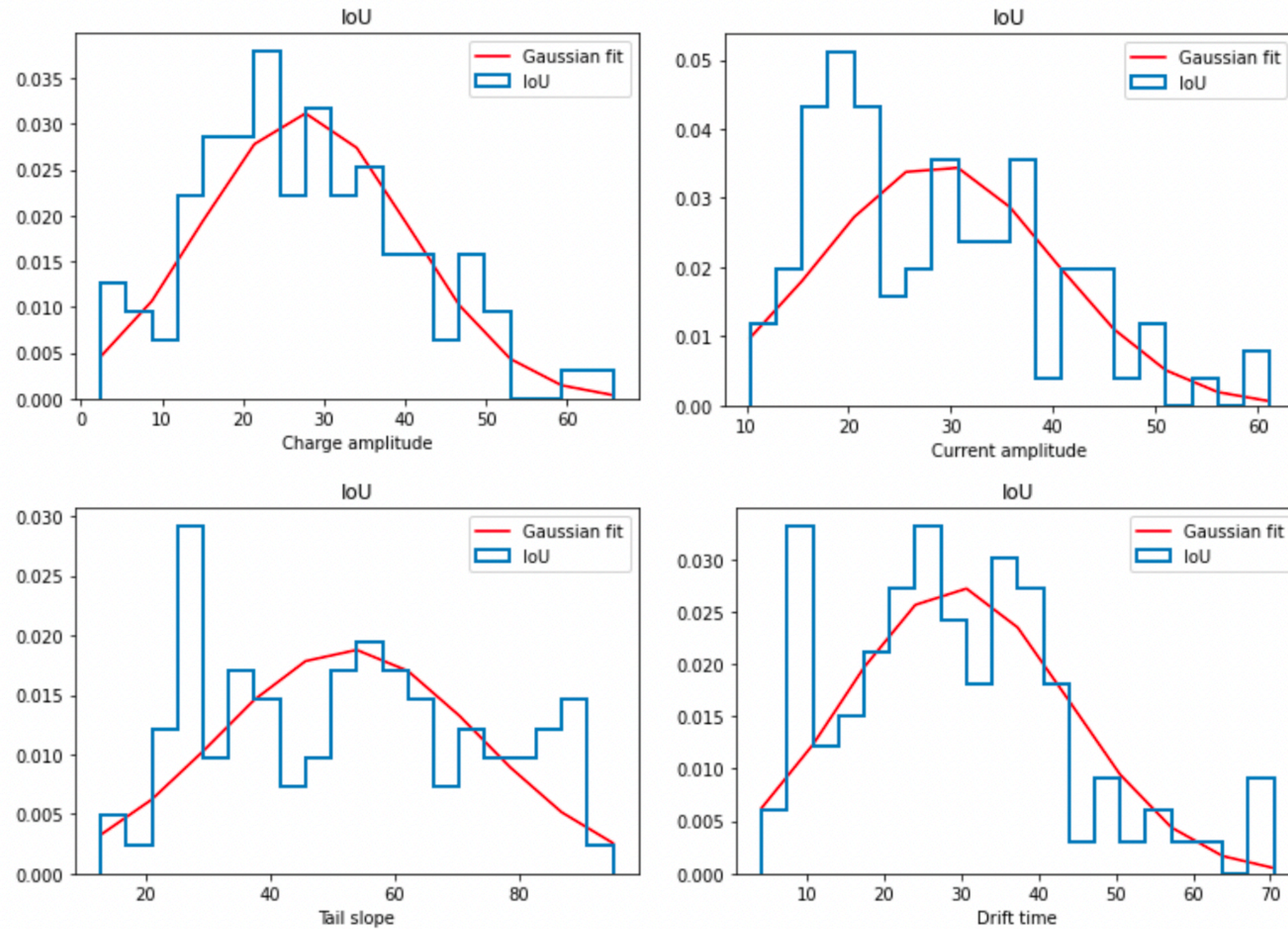


Рисунок 17 - Примеры импульсов, полученных в результате обучения GAN с PCA

Рисунок 18 - Распределения параметров импульсов, полученных в результате обучения GAN с PCA

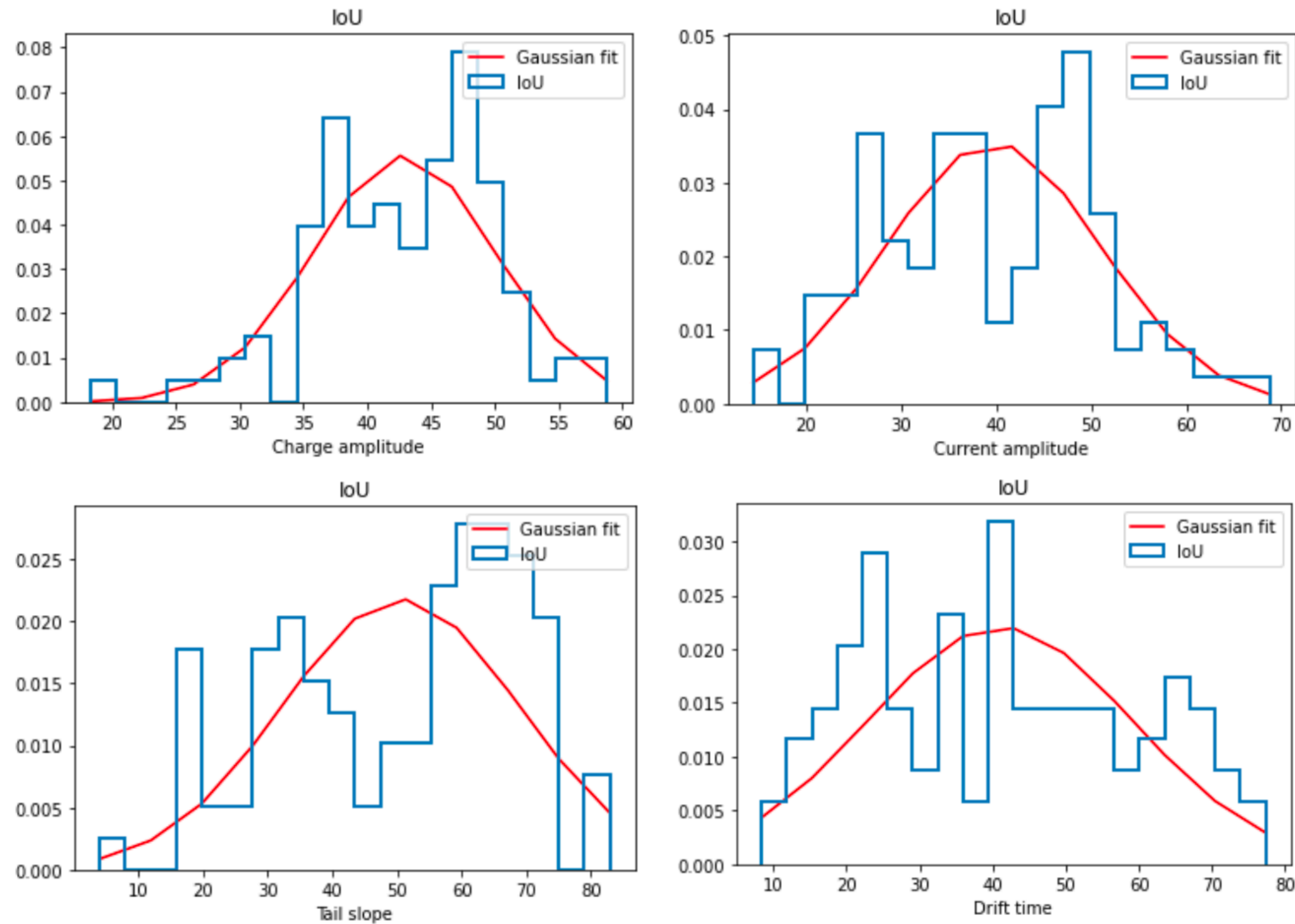
Оценка эффективности модели GAN с использованием PCA



Charge amplitude: IoU = $(27.5 \pm 12.8) \%$
Current amplitude: IoU = $(28.6 \pm 11.4) \%$
Tail slope: IoU = $(52.8 \pm 21.2) \%$
Drift time: IoU = $(29.3 \pm 14.6) \%$

Рисунок 19 - Распределения IoU для модели GAN с PCA,
обученной на данных HADES

Оценка эффективности модели GAN с использованием PCA



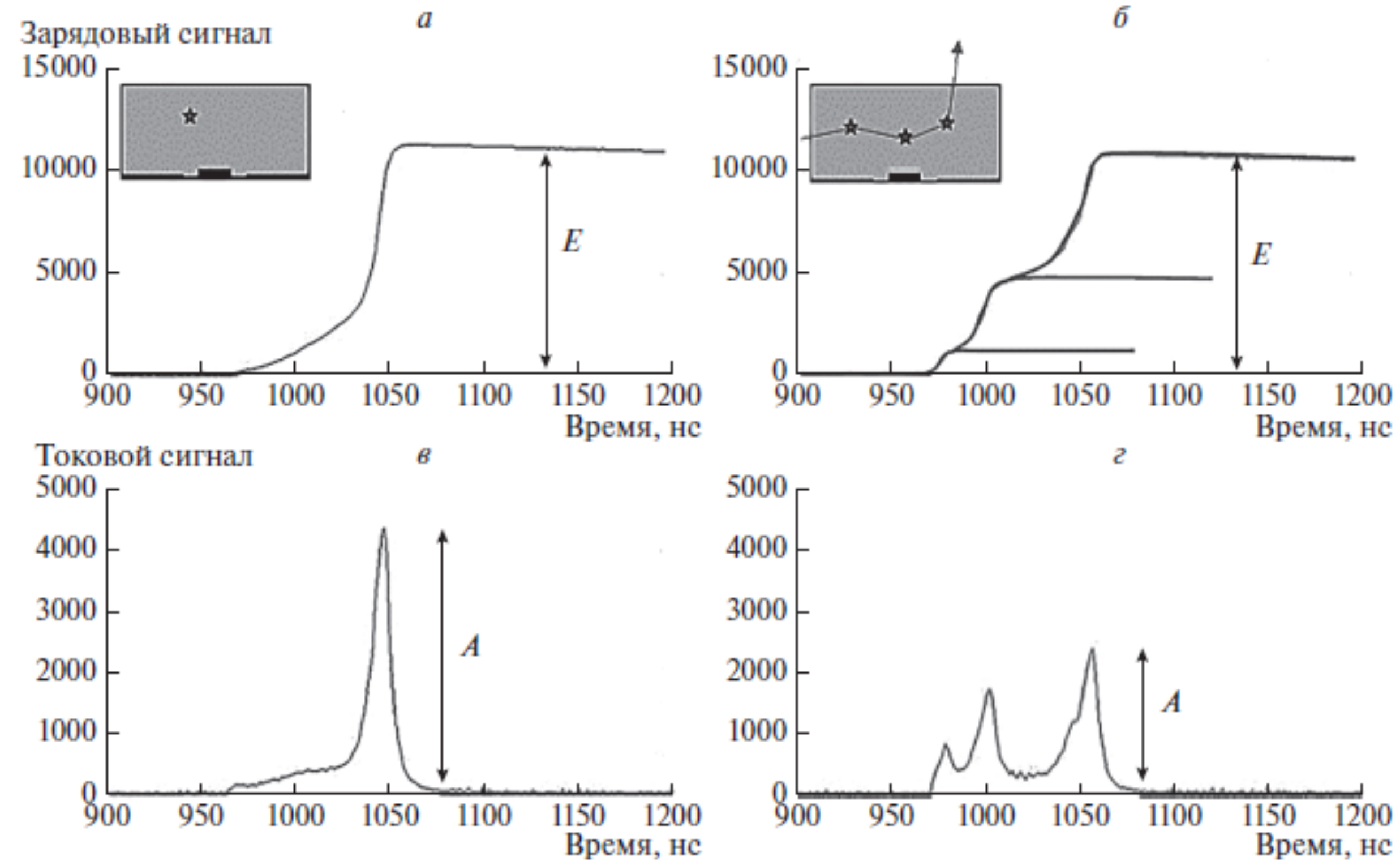
Charge amplitude: IoU = $(42.9 \pm 7.2) \%$
Current amplitude: IoU = $(39.8 \pm 11.3) \%$
Tail slope: IoU = $(50.5 \pm 18.3) \%$
Drift time: IoU = $(41.0 \pm 18.0) \%$

Рисунок 20 - Распределения IoU для модели GAN с PCA,
обученной на данных ОИЯИ

Результаты

- Наиболее приемлемый результат был достигнут в модели GAN с применением метода главных компонент (PCA).
- Модель GAN-PCA была успешно обобщена на данные, полученные на детекторе Sp5020 ORTEC.
- Применение критериев отбора для синтезированных данных необходимо.

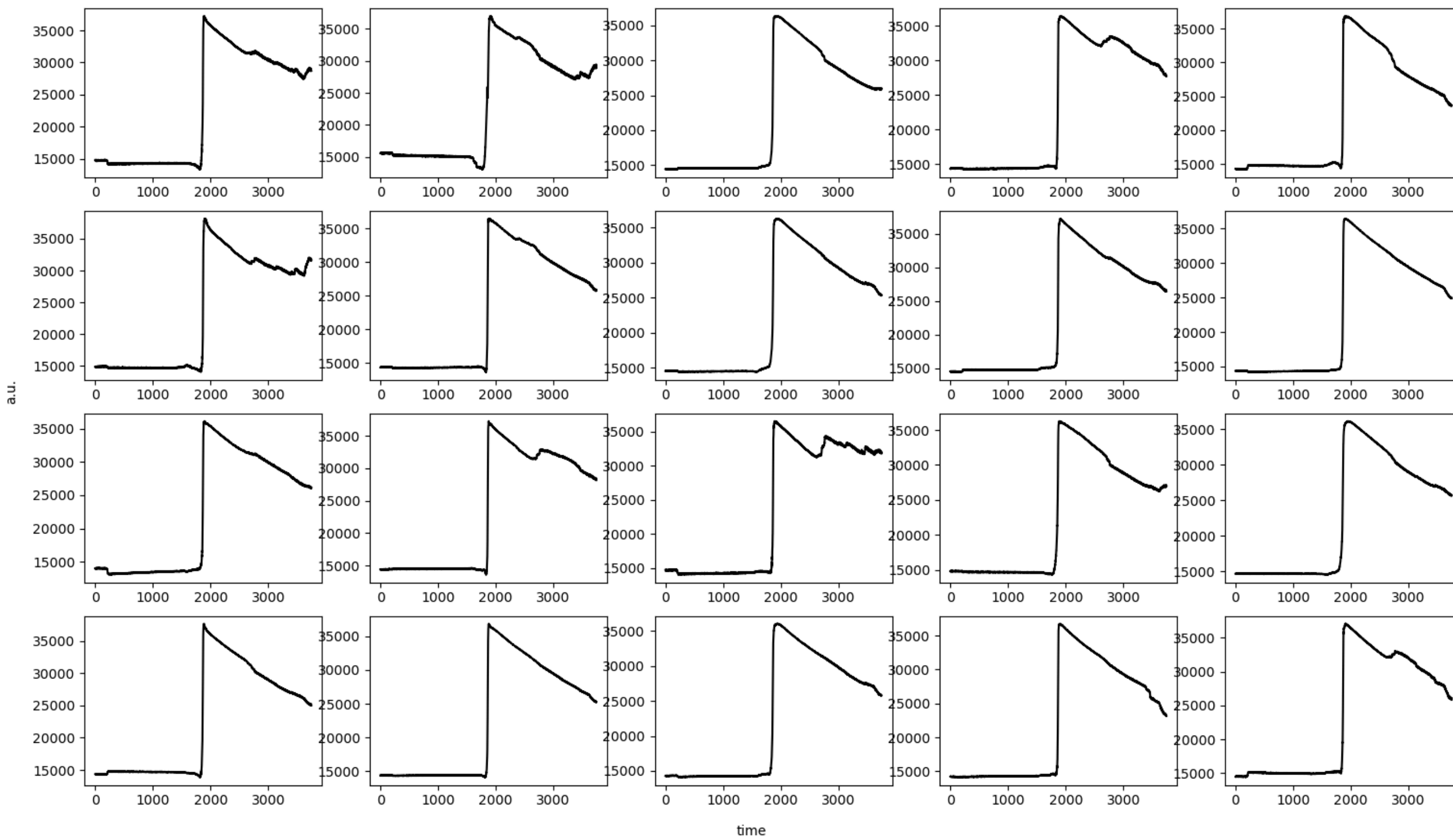
Back-up



Определение параметра отбора A/E , полученного с учетом различия между сигналами SSE и MSE

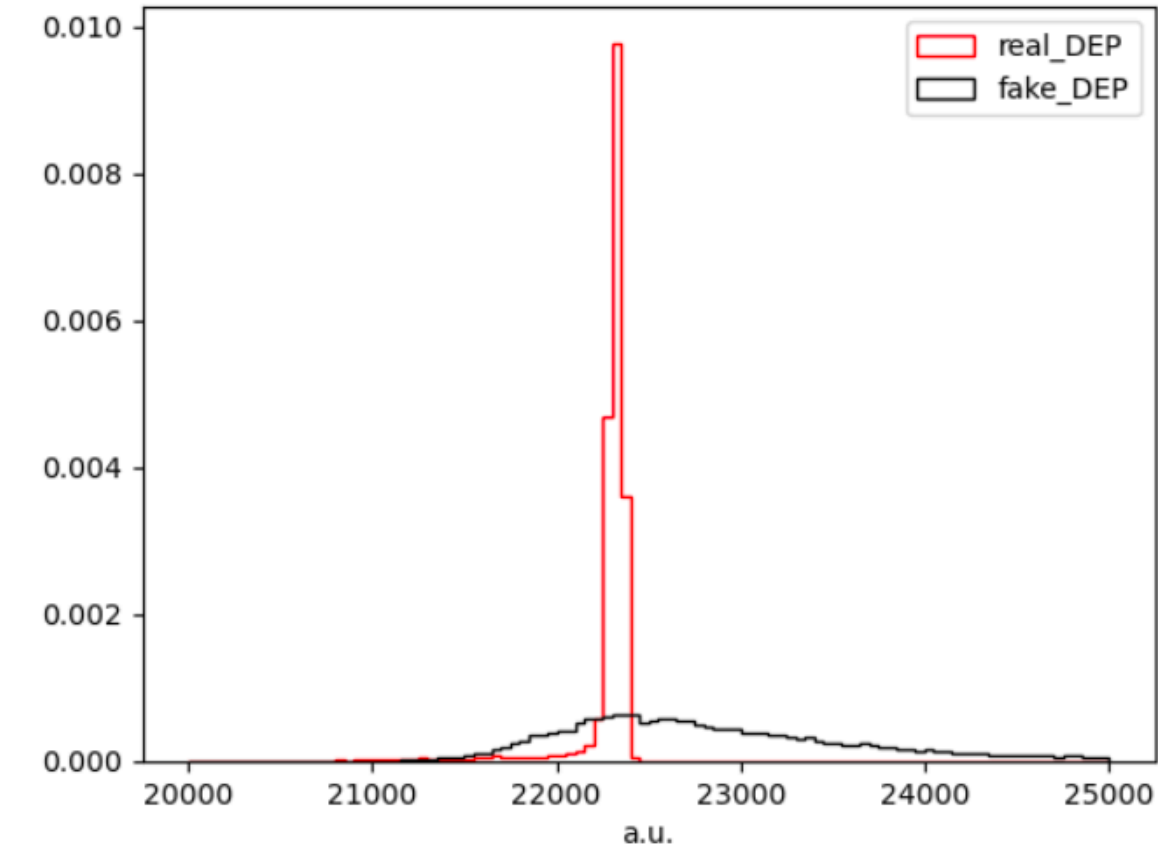
Back-up

Fake DEP

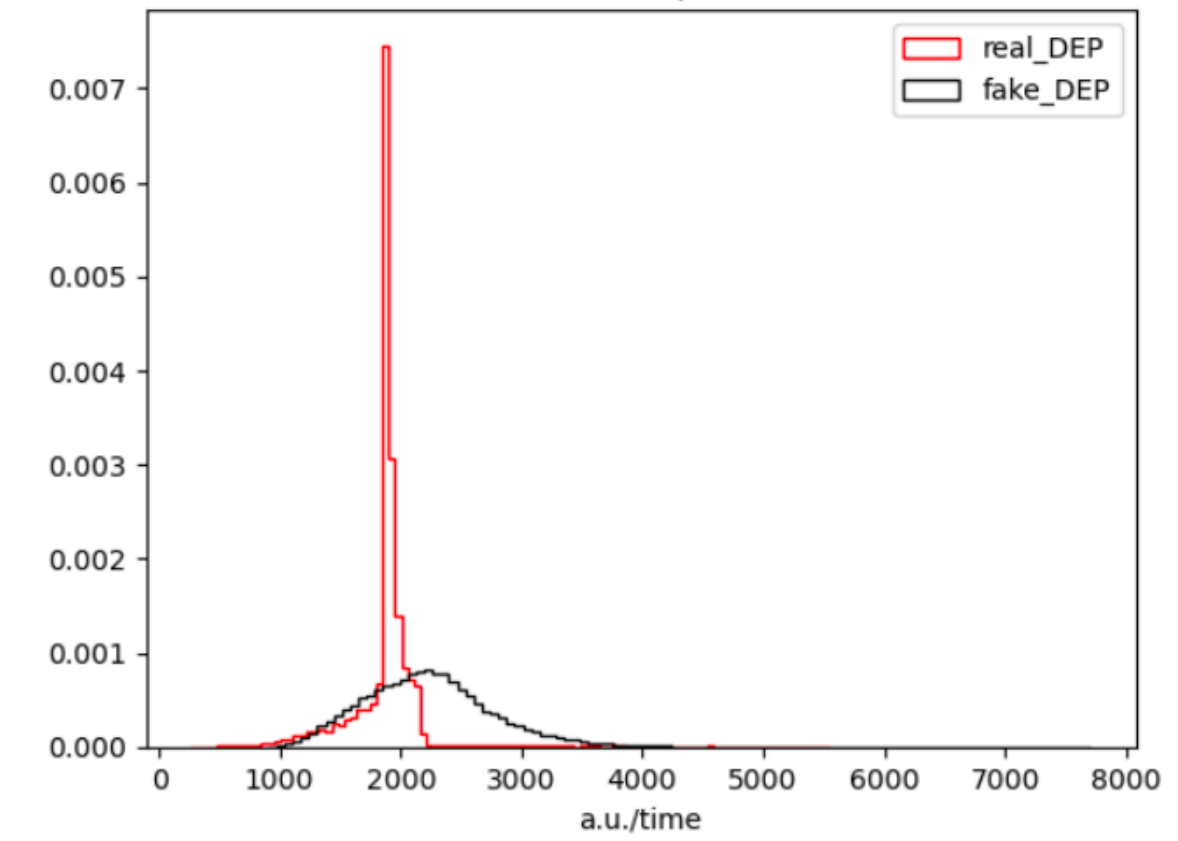


Примеры импульсов, полученных в результате обучения модели GAN с использованием автокодировщика

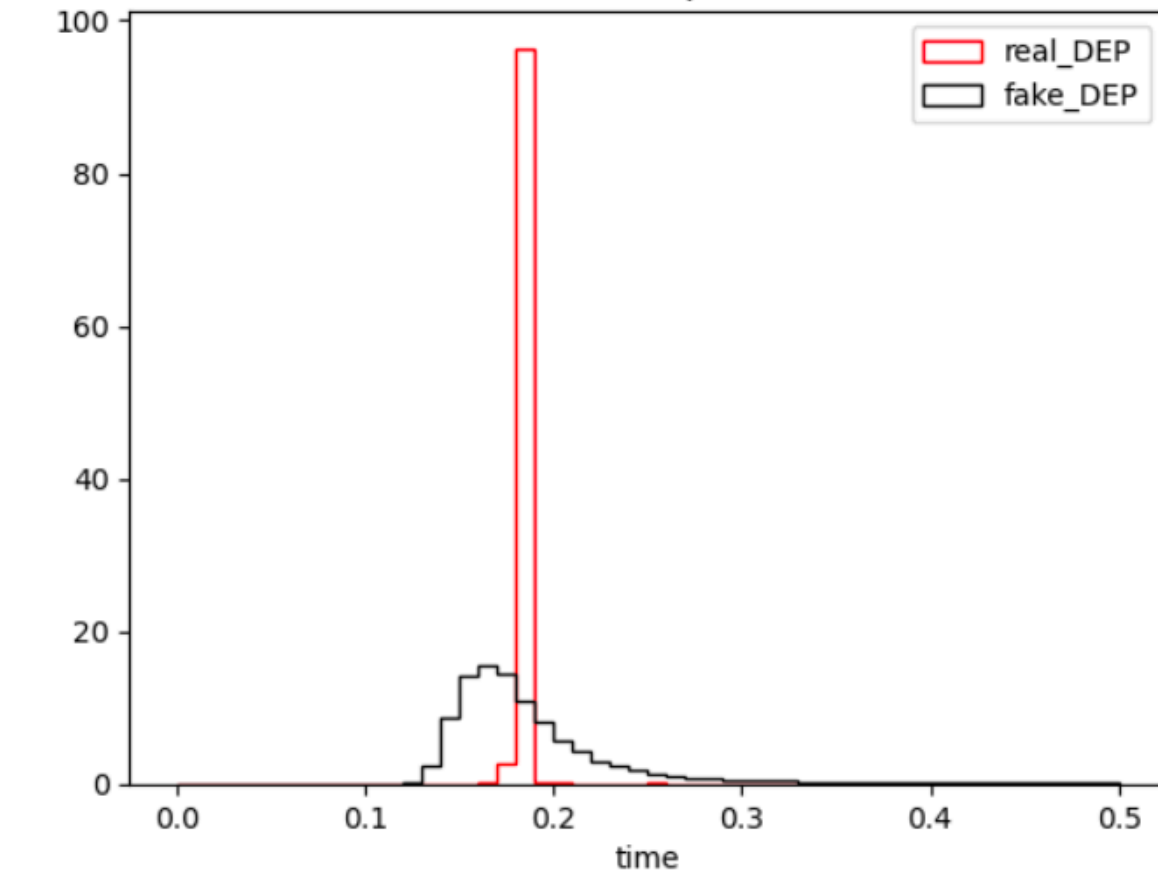
Charge amplitude



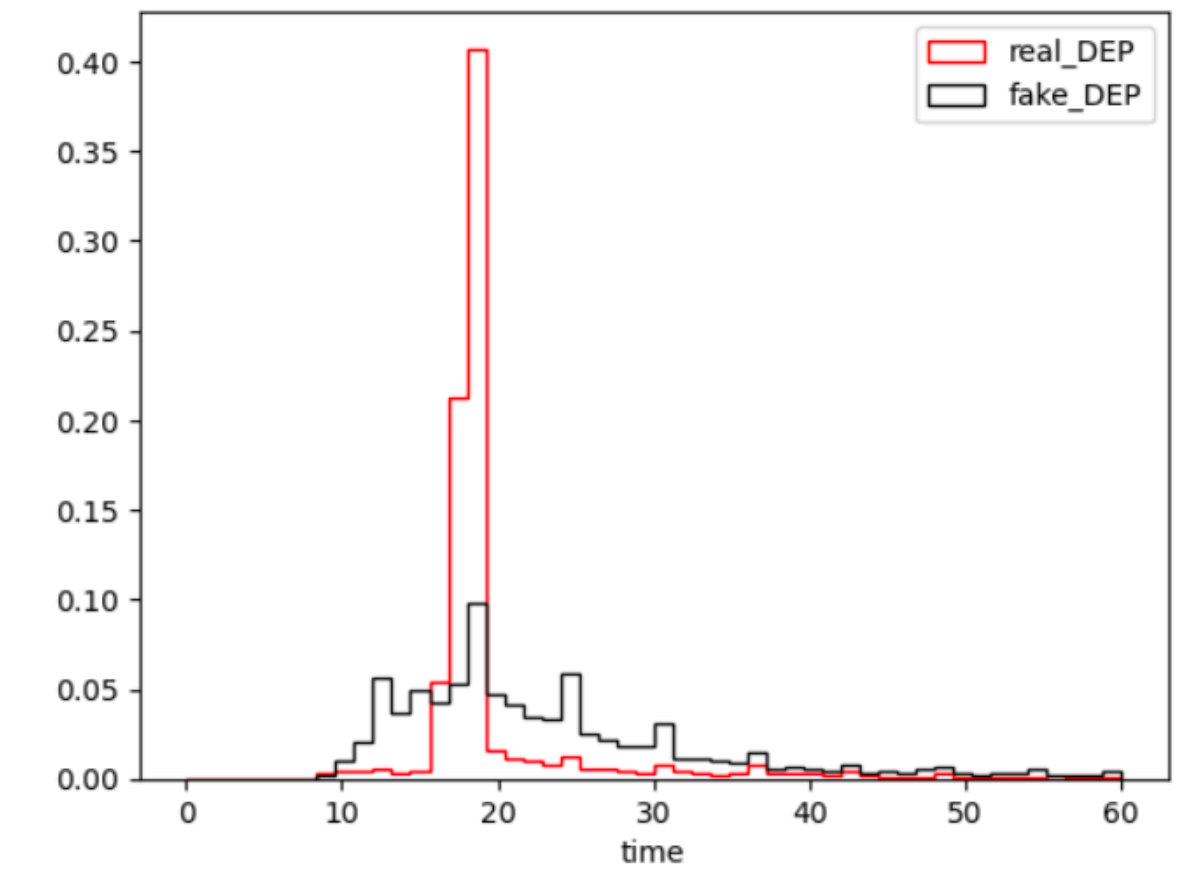
Current amplitude



Tail slope



Drift time



Распределения параметров импульсов, полученных в результате обучения модели GAN с использованием автокодировщика

Вак-уп

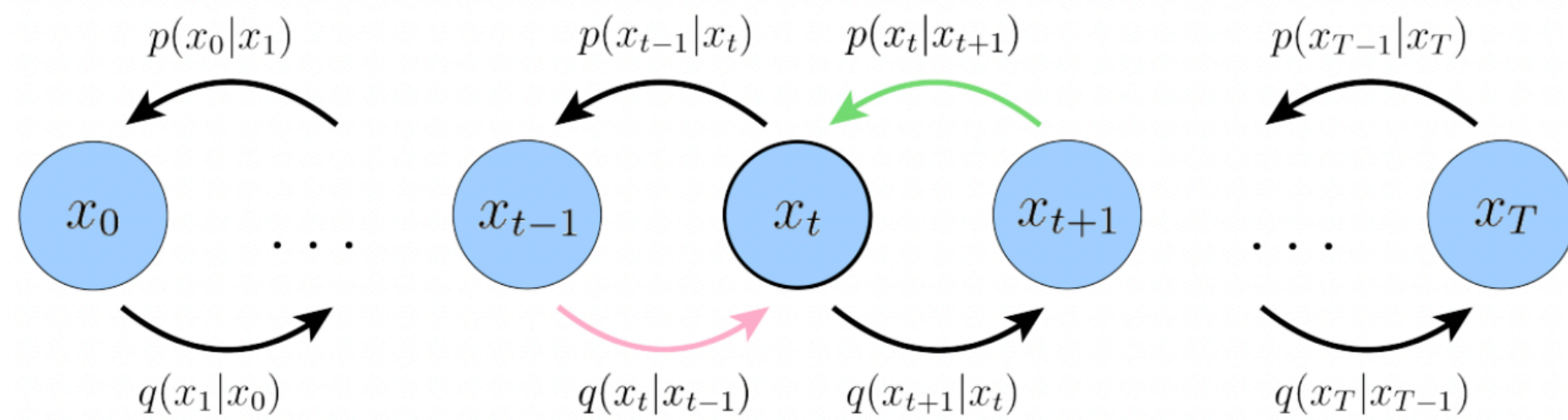
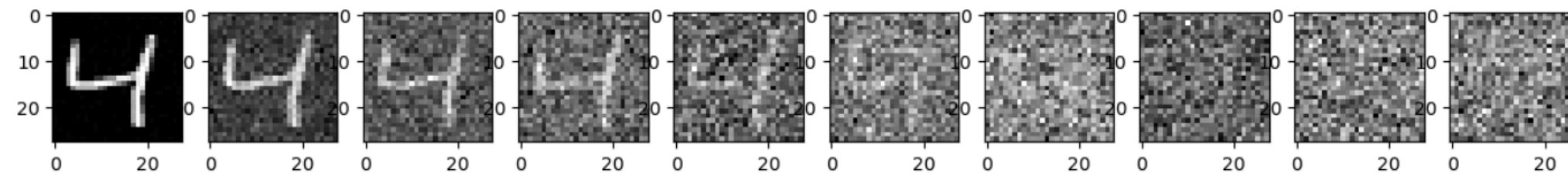
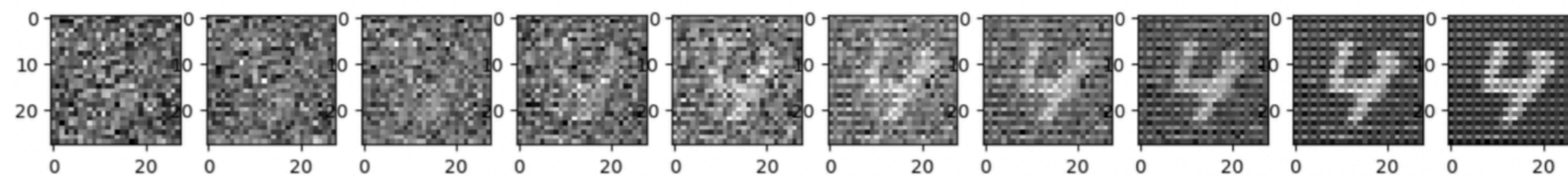


Схема работы диффузионной модели



Прямой процесс

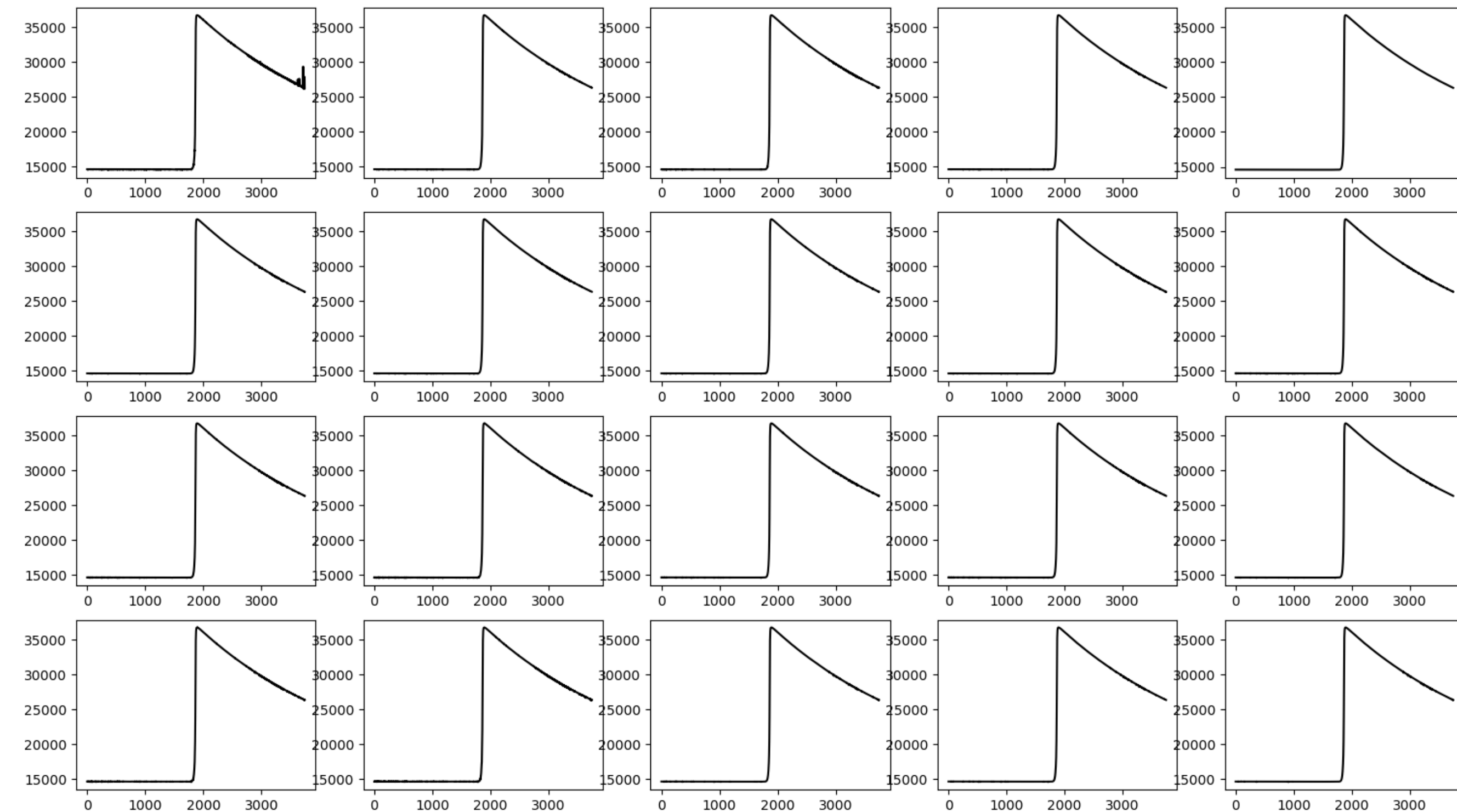


Обратный процесс

- **Прямой процесс:** $x_t = \sqrt{\bar{\alpha}_t} \cdot x_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \epsilon$, где $\bar{\alpha}_t$ - заранее выбранные коэффициенты
 ϵ - случайный шум
- Для **обратного процесса** требуется нейронная сеть, предсказывающая уровень «зашумленности» данных на текущем уровне: $x_{t-1} \approx x_t - noise$
- Коэффициенты $\bar{\alpha}_t$ выбираются так, чтобы на последнем уровне (T) получался случайный шум
- Процесс генерации данных, в отличие от модели GAN, итеративный

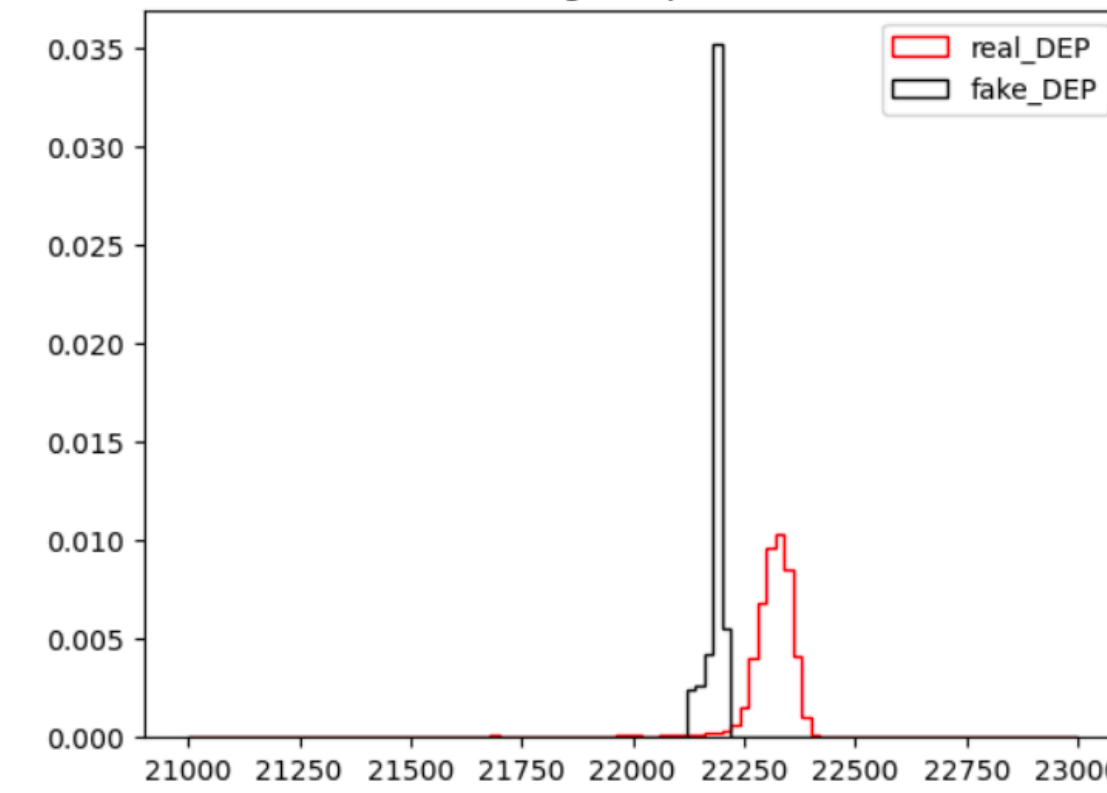
Back-up

Fake DEP

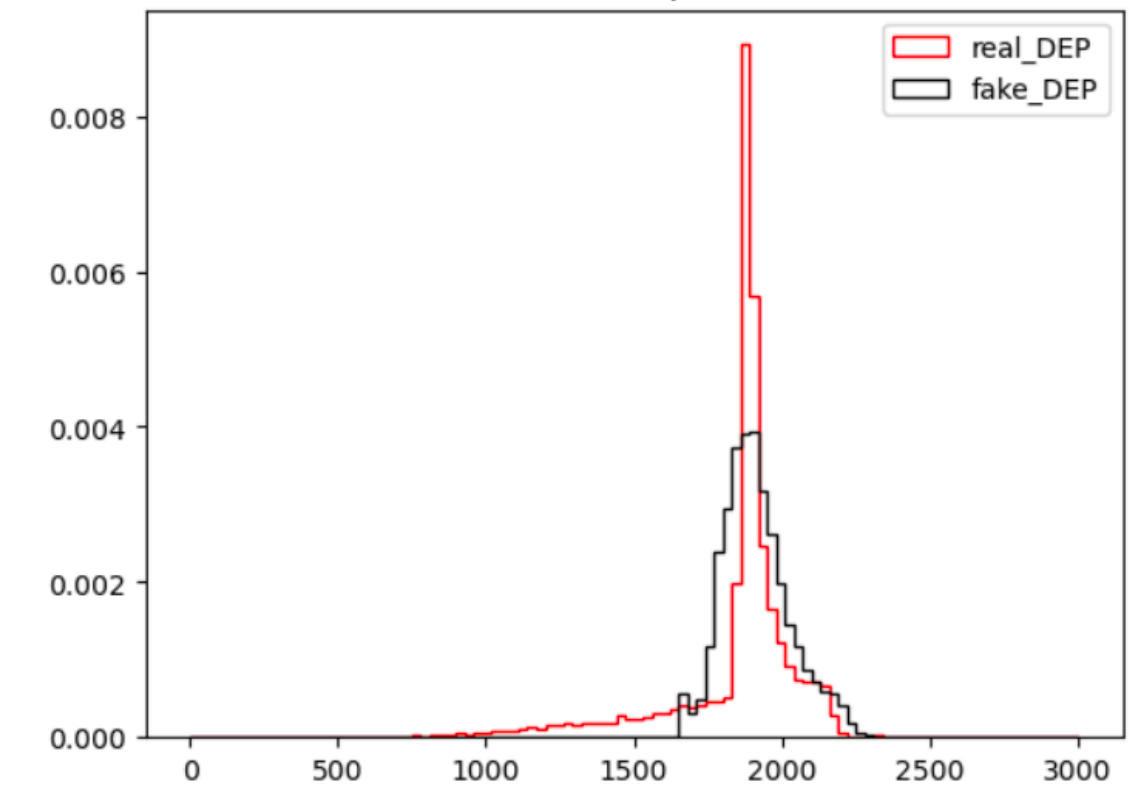


Примеры импульсов, полученных в результате обучения диффузионной модели

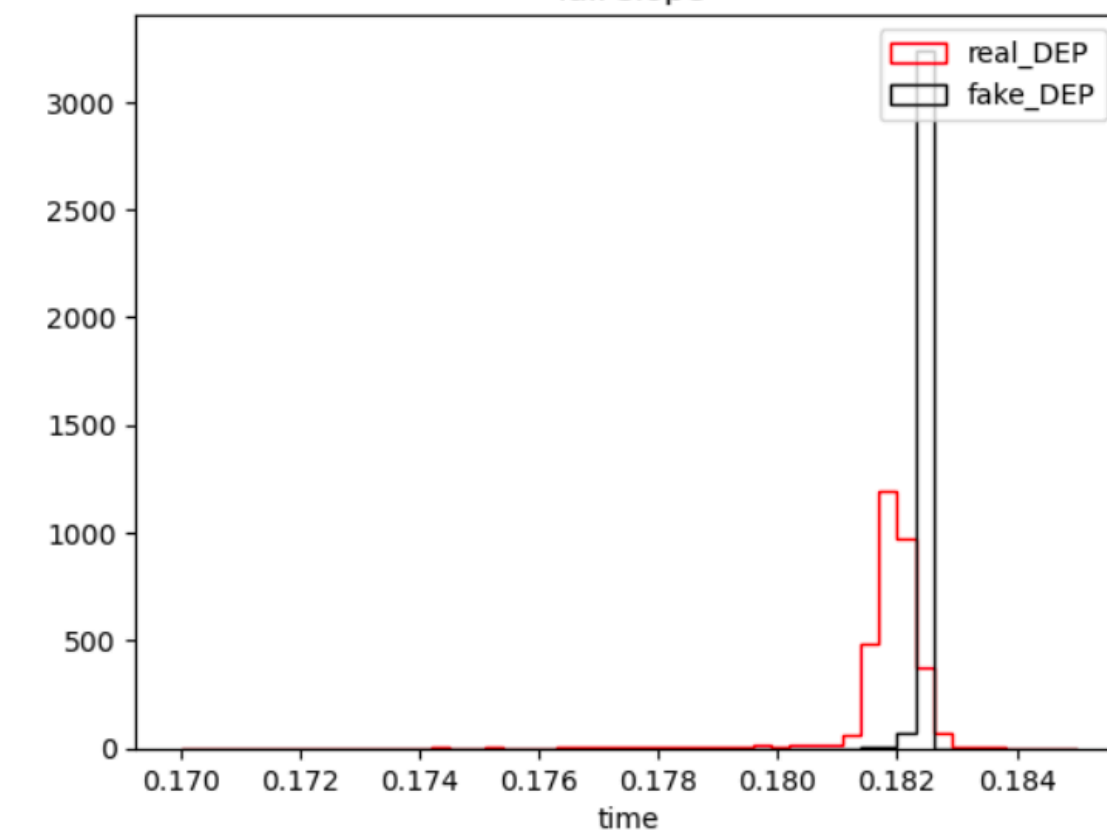
Charge amplitude



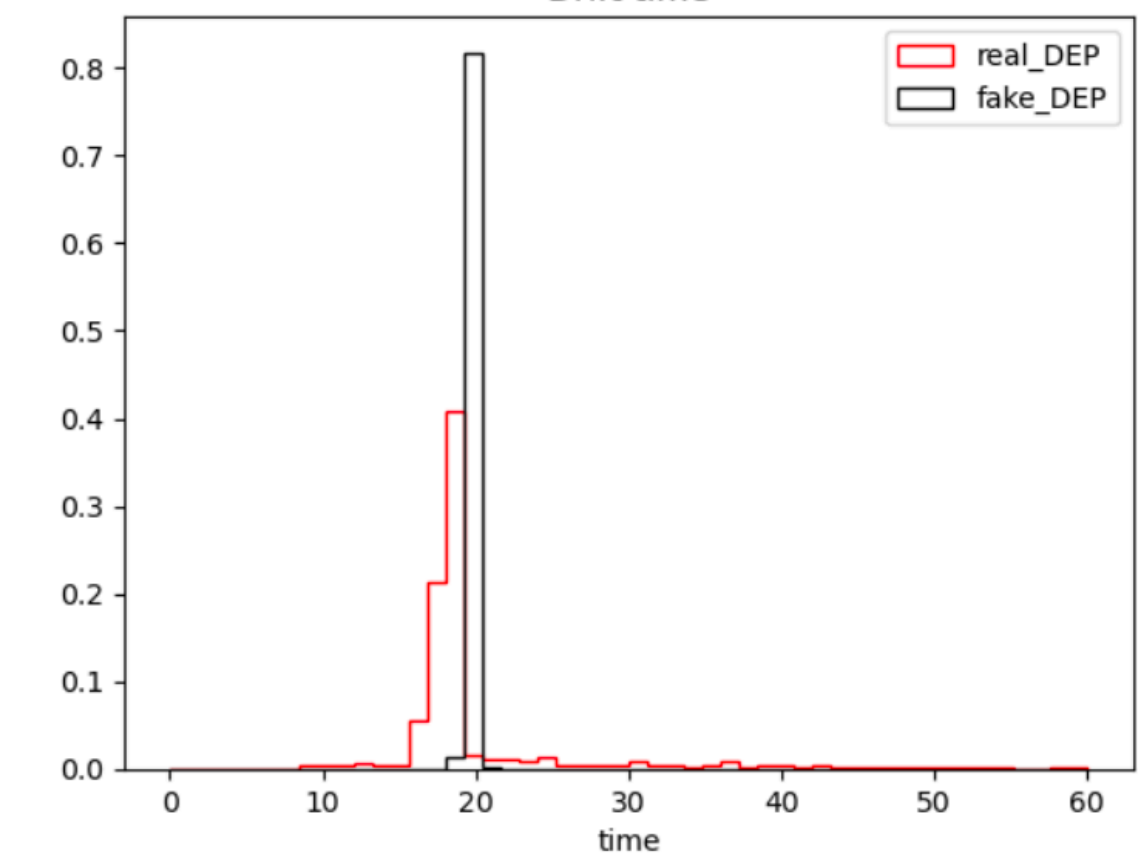
Current amplitude



Tail slope

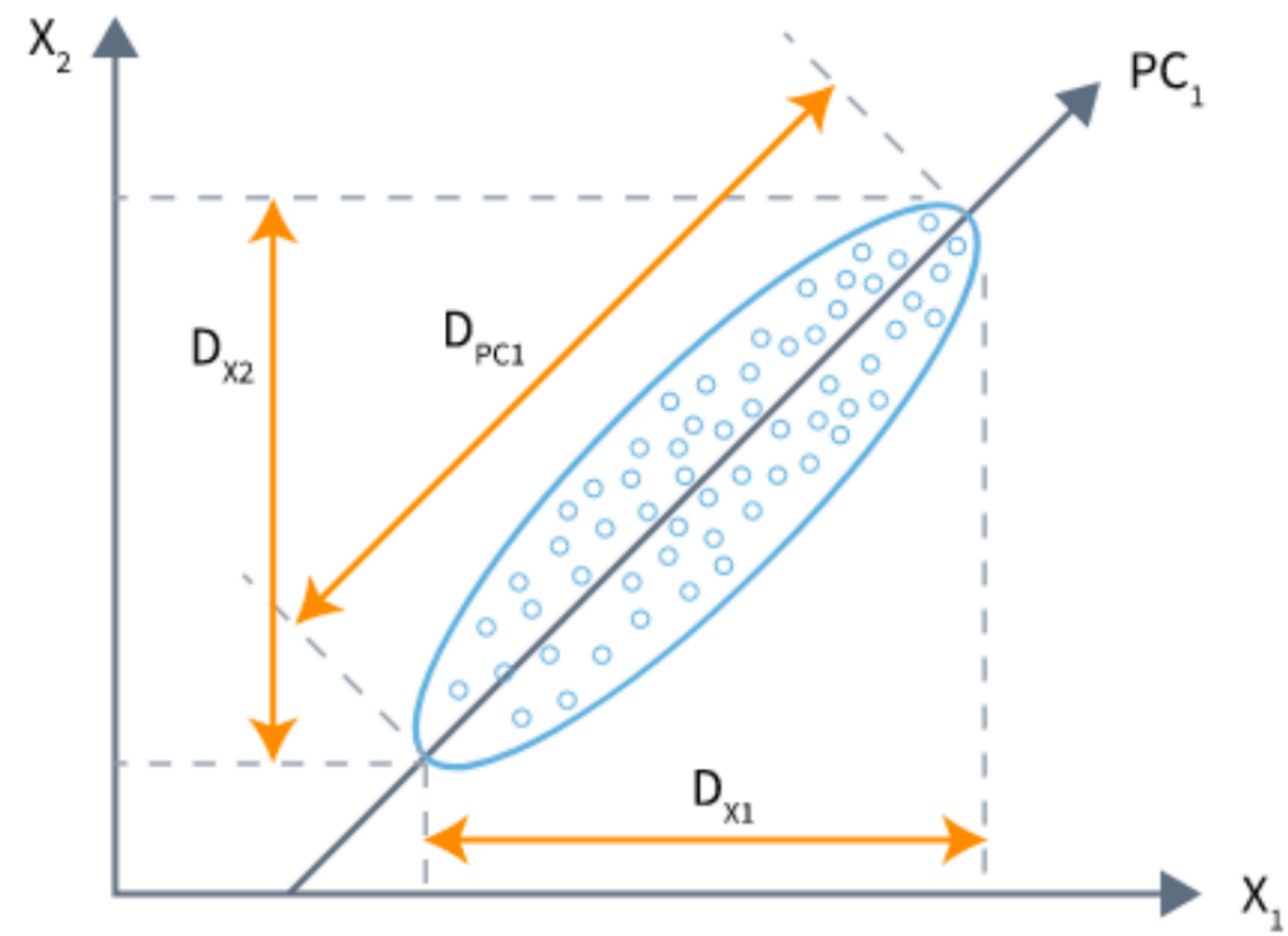


Drift time



Распределения параметров импульсов, полученных в результате обучения диффузионной модели

Back-up



Снижение размерности исходного
2-мерного пространства до 1-мерного