Подавление шумовых хитов и событий от ШАЛ нейронными сетями с применением доменной адаптации в эксперименте Baikal-GVD

Работа выполнена при поддержке гранта РНФ № 24-72-10056.

Мацейко Альберт
МФТИ, ИЯИ РАН
matseiko.av@phystech.edu

Март 2026

План

1. Телескоп ***Baikal-GVD***: устройство и данные

2. ***Схема решения задач*** эксперимента нейронными сетями

3. Результаты ***подавления фона от ШАЛ***:
   2.1 На Монте-Карло симуляции
   2.2 Для реальных данных: польза Domain Adaptation (DA)

4. Результаты ***подавления шумовых хитов***:
   3.1 На Монте-Карло симуляции
   3.2 Для реальных данных (с применением DA)

Результаты коллеги
Григорий Плотников,
Plotnikov.gp@phystech.edu

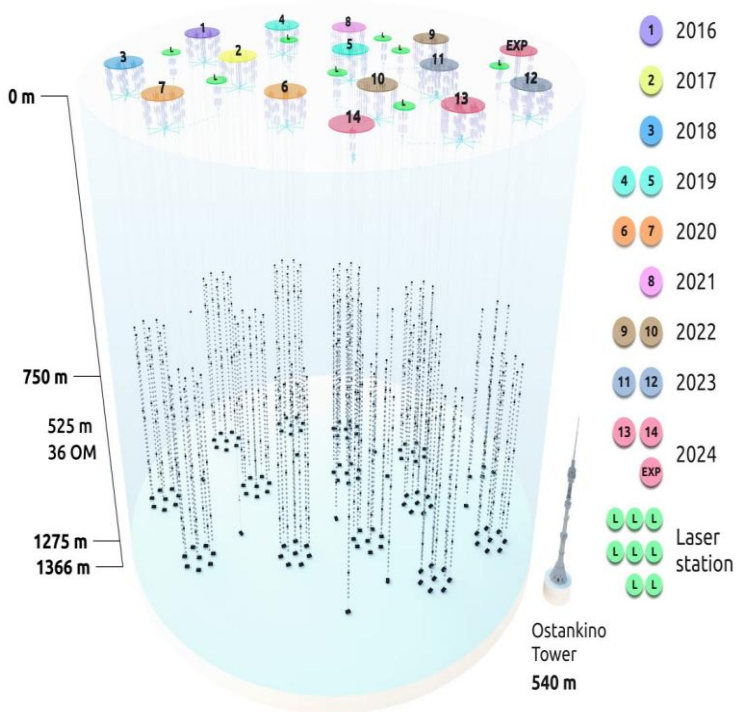
0. Как устроен Baikal-GVD?

Оптические Модули (ФЭУ)

Собираются в «струны»

→

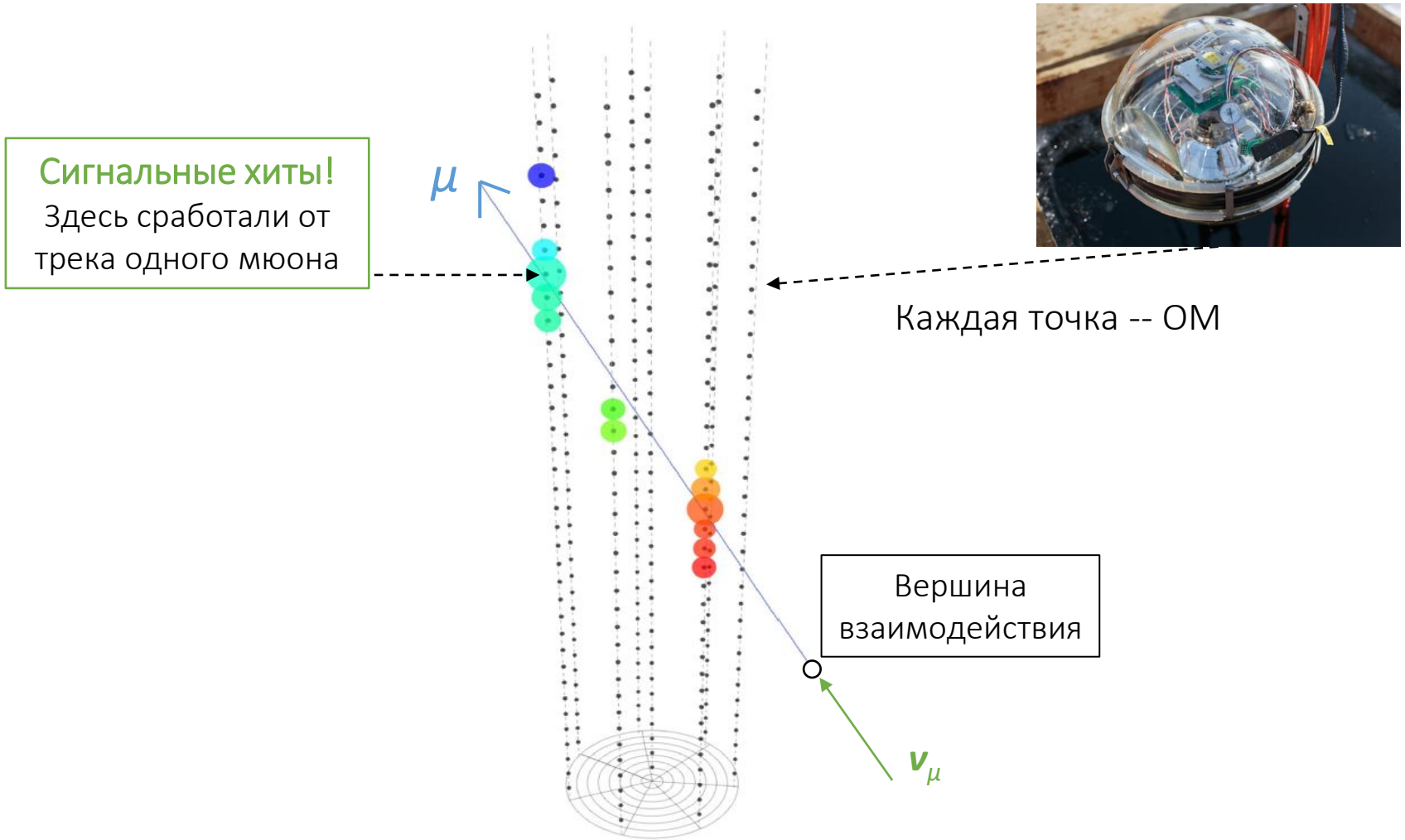
собираются в «кластеры»



Установлено >14 кластеров

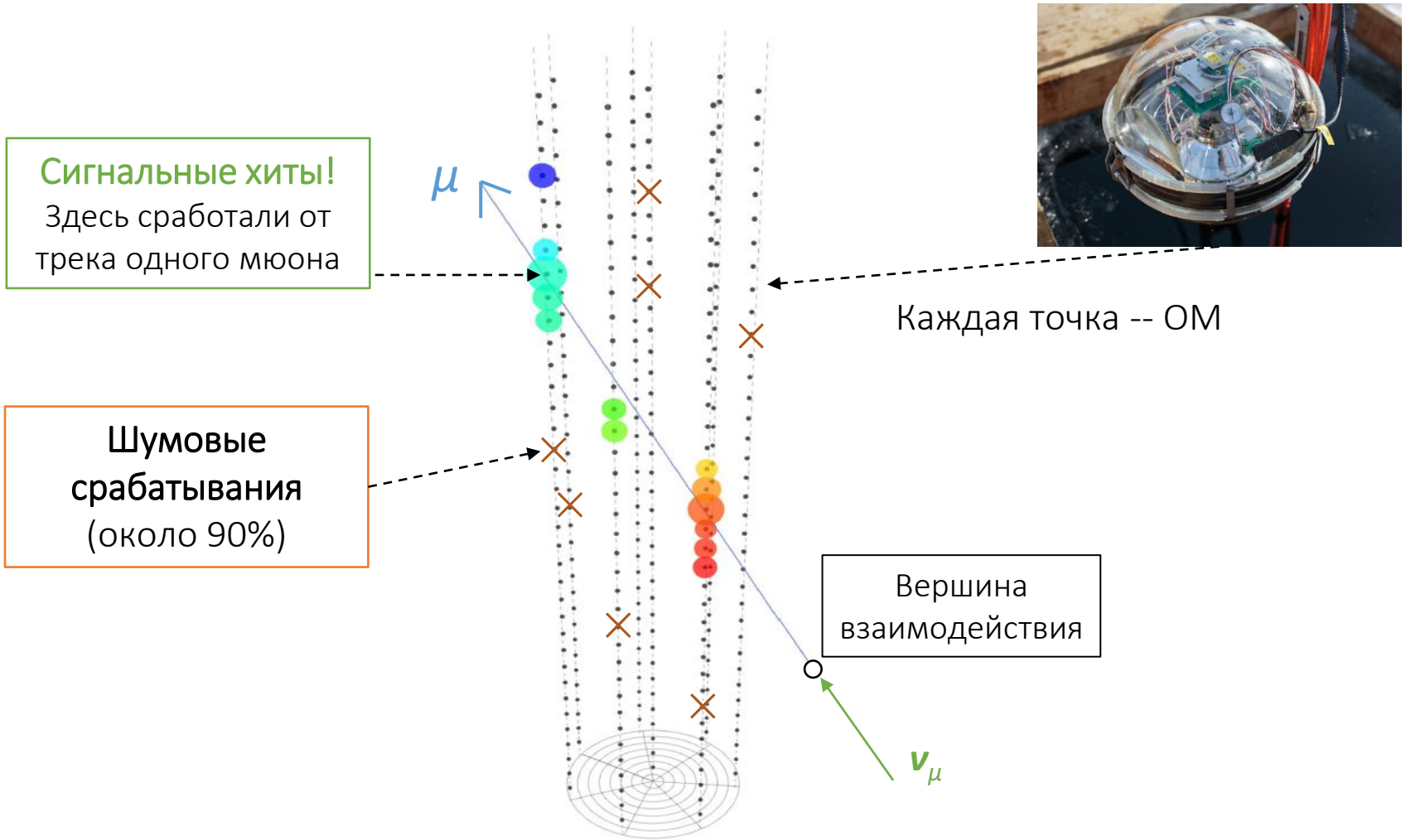
Процесс установки

1. Данные в Baikal-GVD: как устроены события?



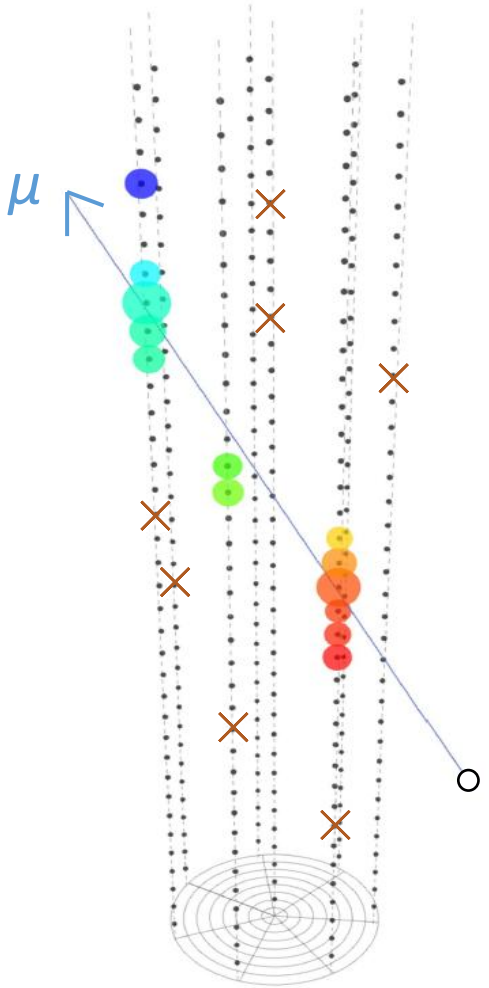
Пример события в пределах одного кластера

1. Данные в Baikal-GVD: как устроены события?



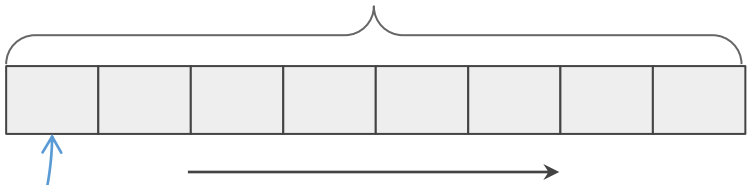
Пример события в пределах одного кластера

1. Данные в Baikal-GVD: как «скормить» нейросети?



Пример события в пределах одного кластера

Последовательность (хитов)

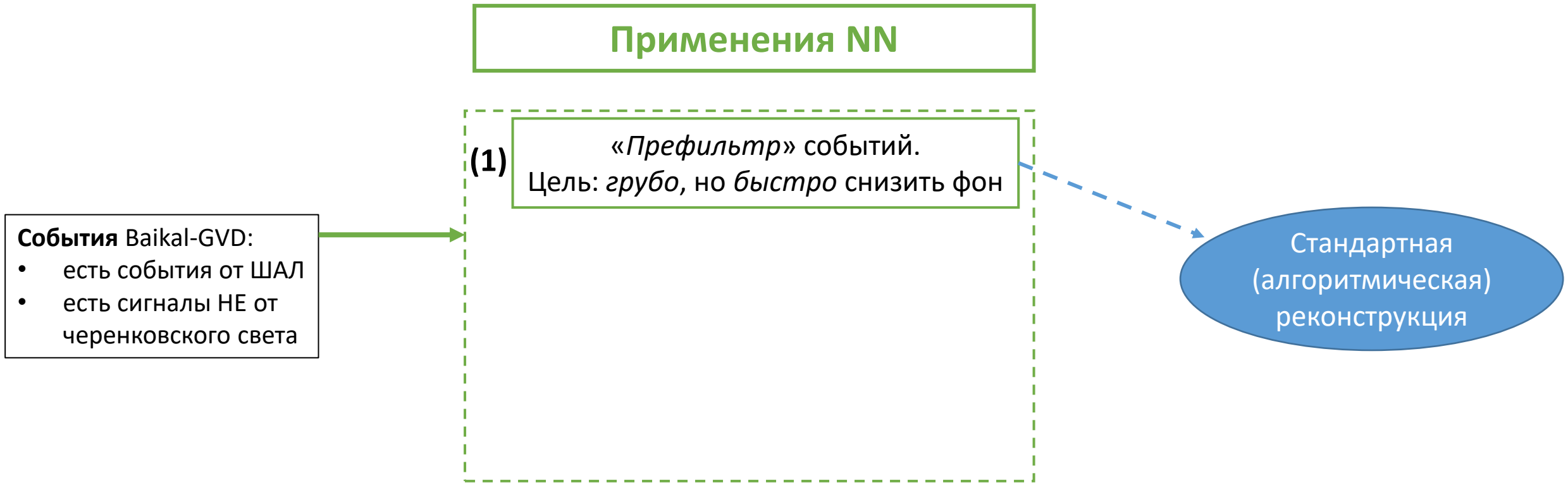


Упорядочиваем по t

1 **хит**
Q, t, x, y, z

Нейронная сеть: *трансформер*
Использует механизм *self-attention*
-->
Ищет важные связи между отдельными хитами
-->
Выявляет сложные зависимости

2. Схема применения нейросетей (NN)



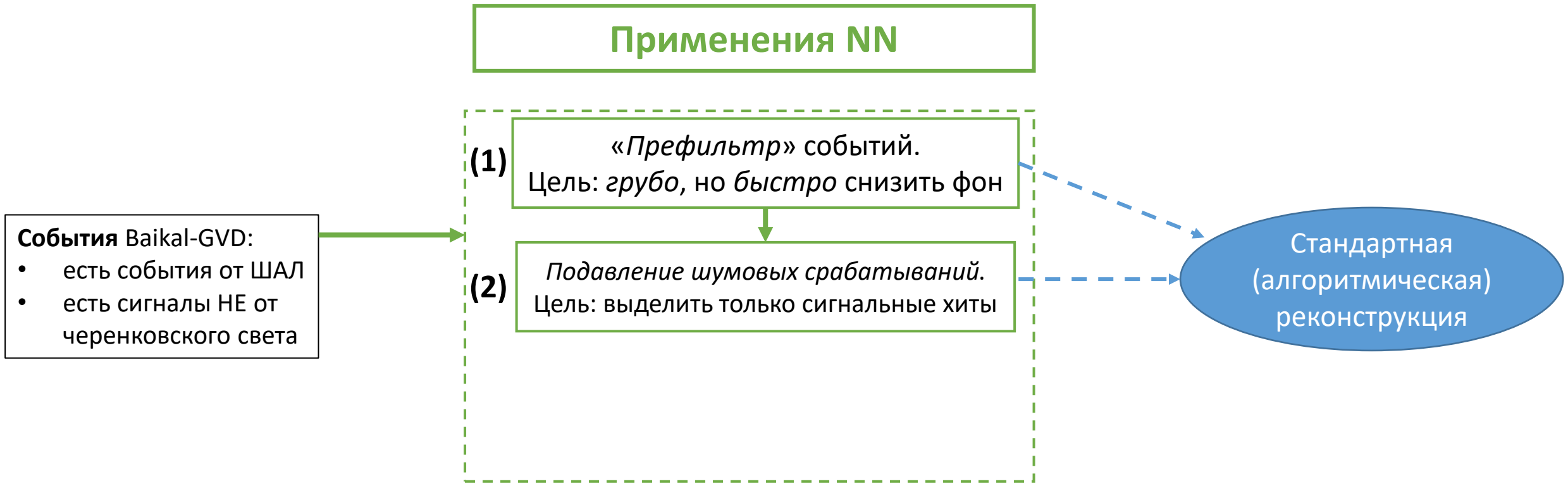
(1)

Проблема: 1 ν на 10^6-10^7 событий.

Решение NN: подавляем ШАЛы в 10 раз, сохраняя 99% событий от ν

Можно ускорить стандартную реконструкцию в разы!

2. Схема применения нейросетей (NN)



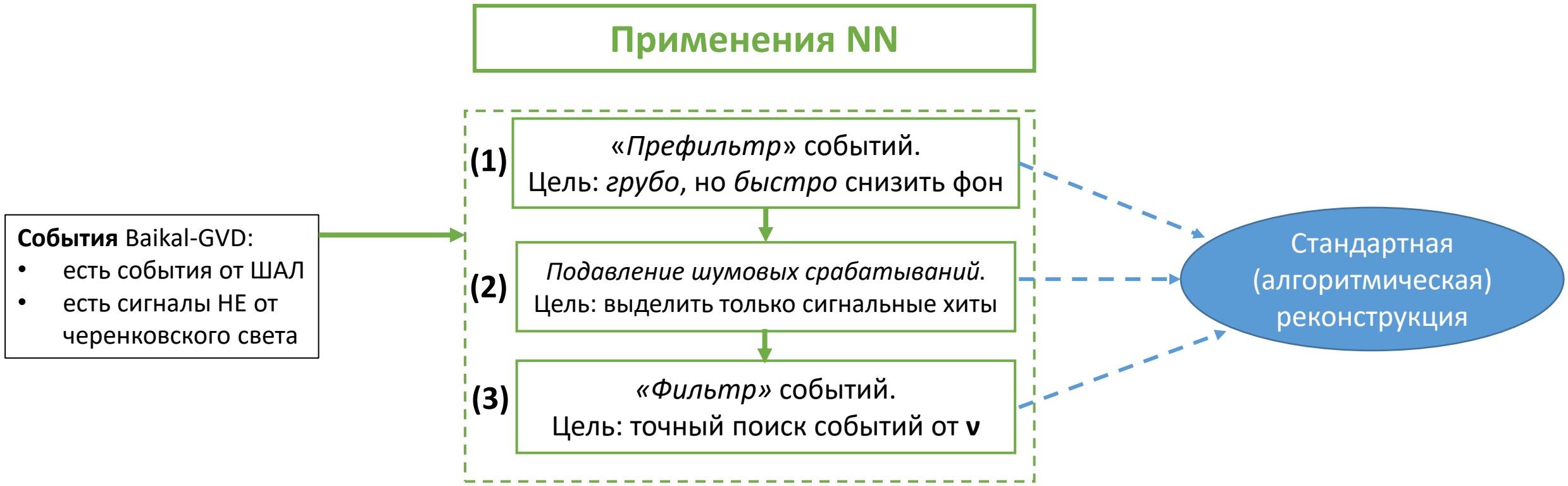
(2)

Проблема: 85%-90% срабатываний ОМ вызваны люминесценцией воды.

Решение NN: подавляем шумовые хиты в >100 раз, сохраняя 95% сигнальных

Улучшает стандартную реконструкцию!

2. Схема применения нейросетей (NN)



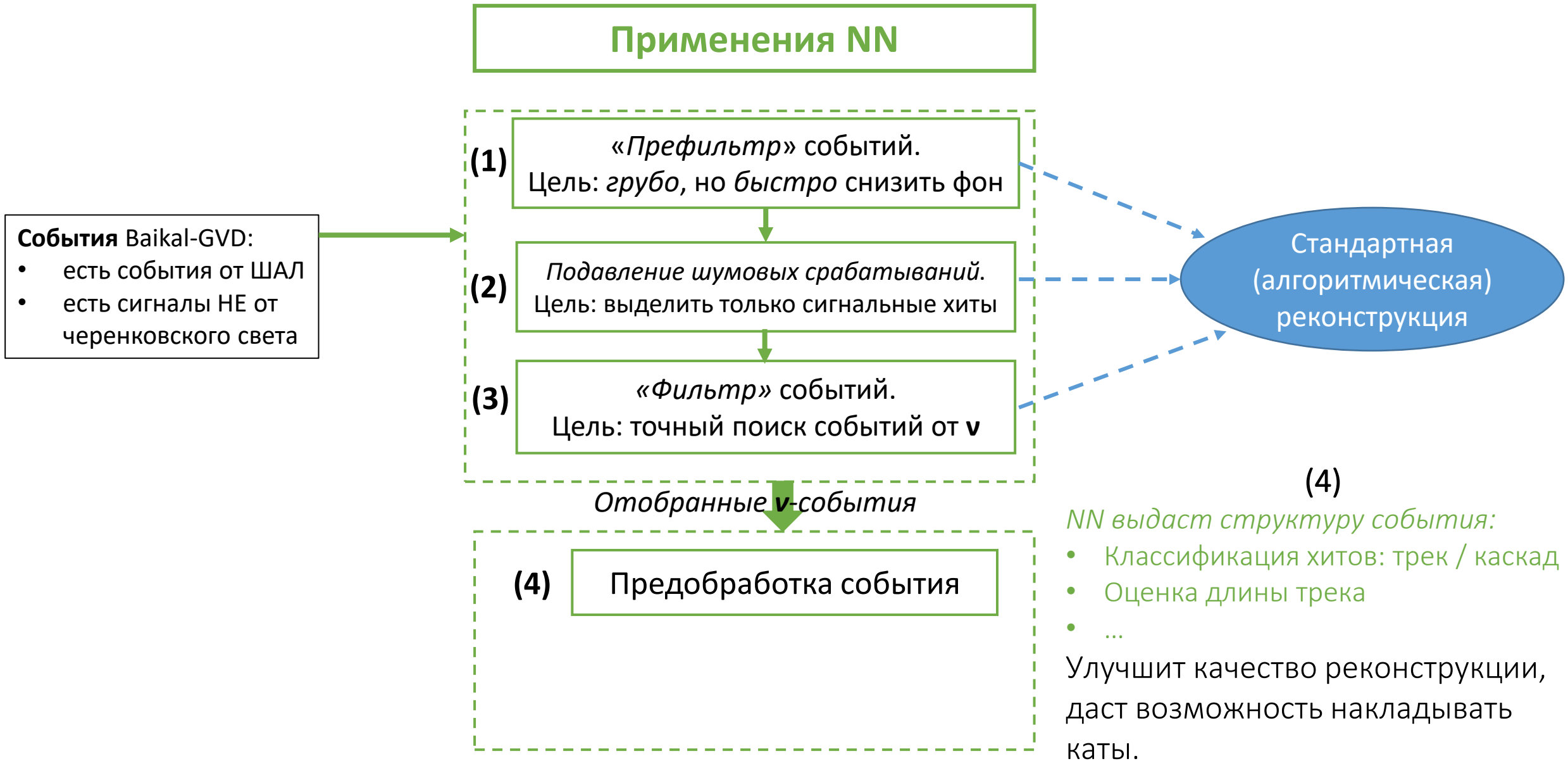
(3)

Проблема: всё ещё много событий от ШАЛ в данных.

Решение NN: позволит составить каталог **ν**-кандидатов и оценить поток!

Найденные события можно подавать и в стандартную реконструкцию!

2. Схема применения нейросетей (NN)



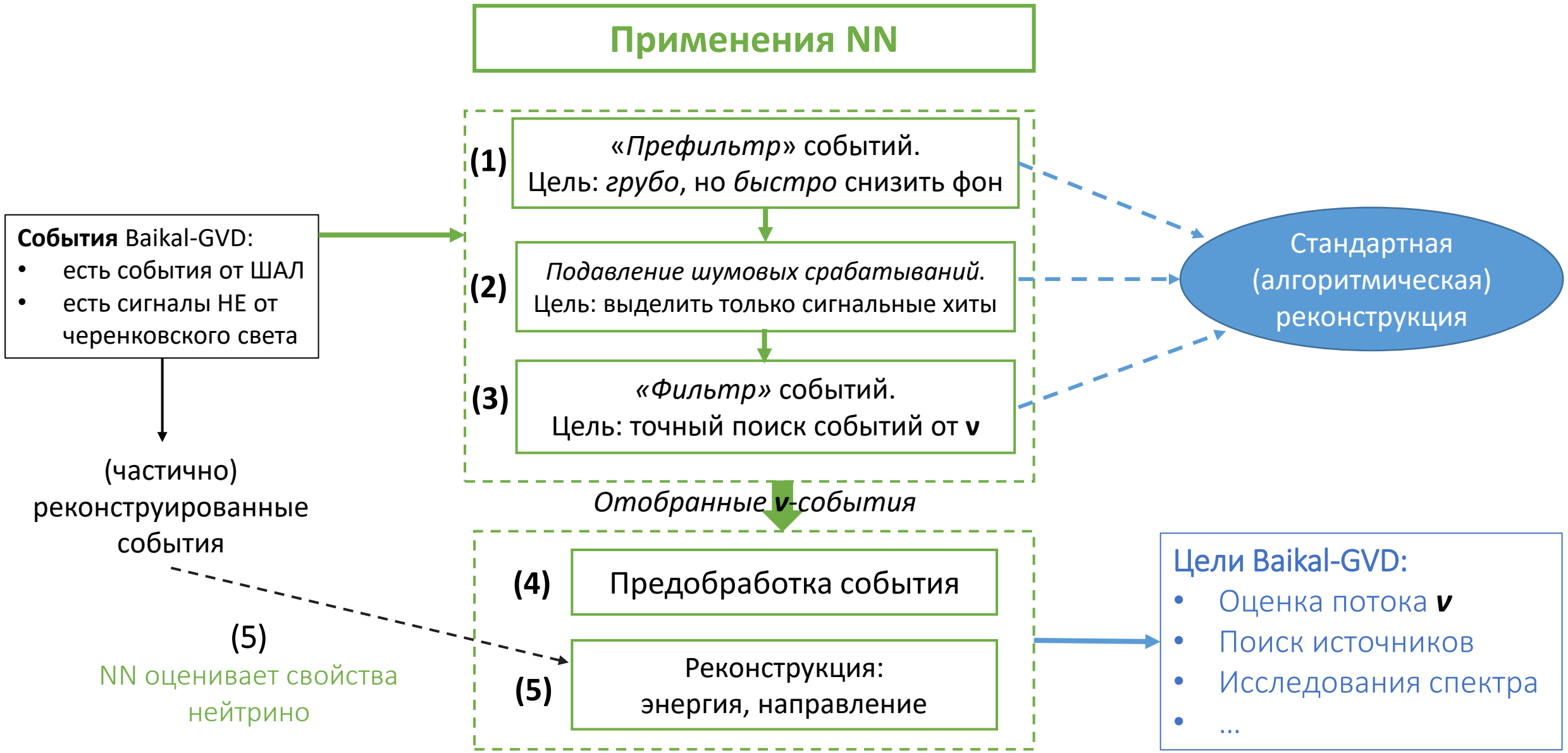
(4)

NN выдаст структуру события:

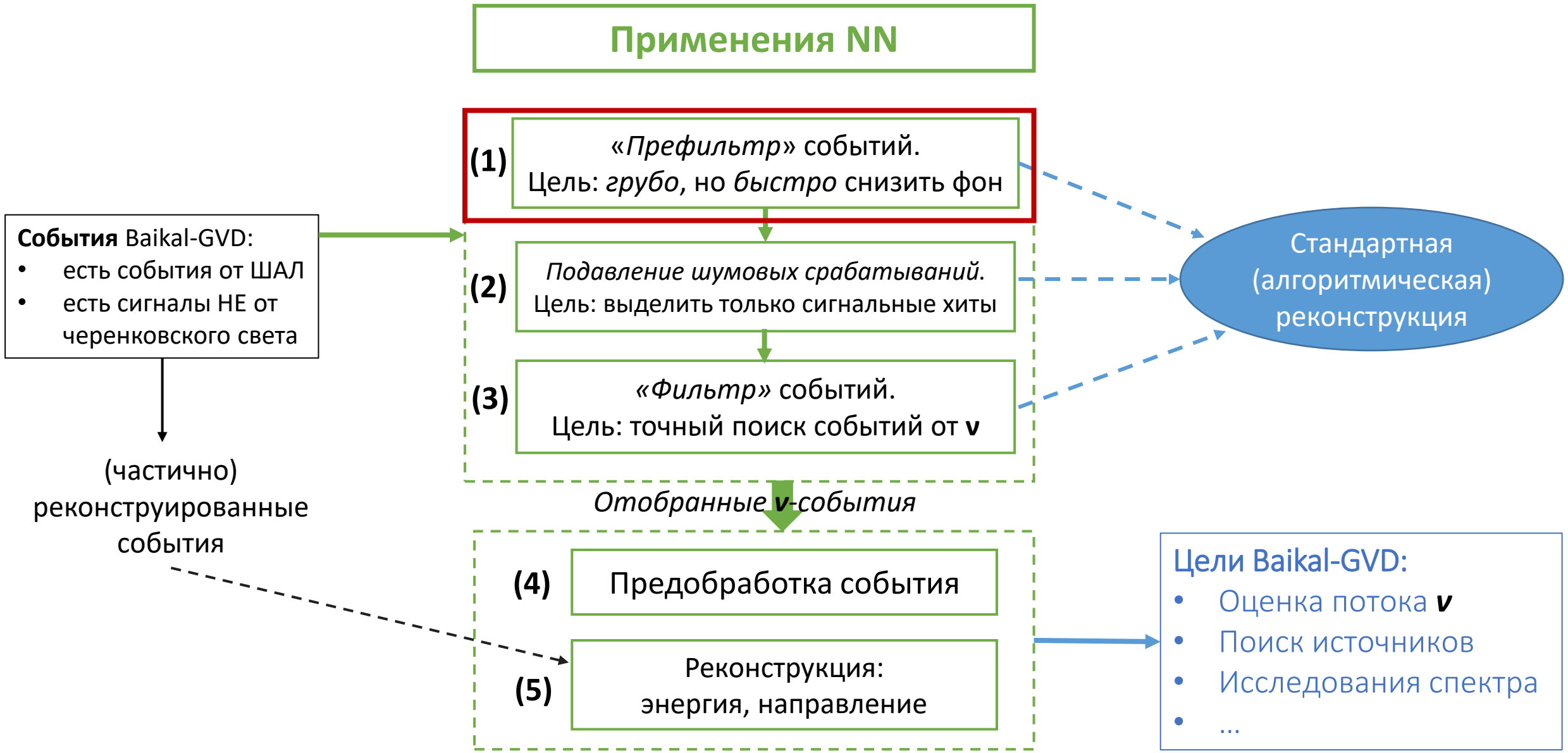
- Классификация хитов: трек / каскад
- Оценка длины трека
- …

Улучшит качество реконструкции, даст возможность накладывать каты.

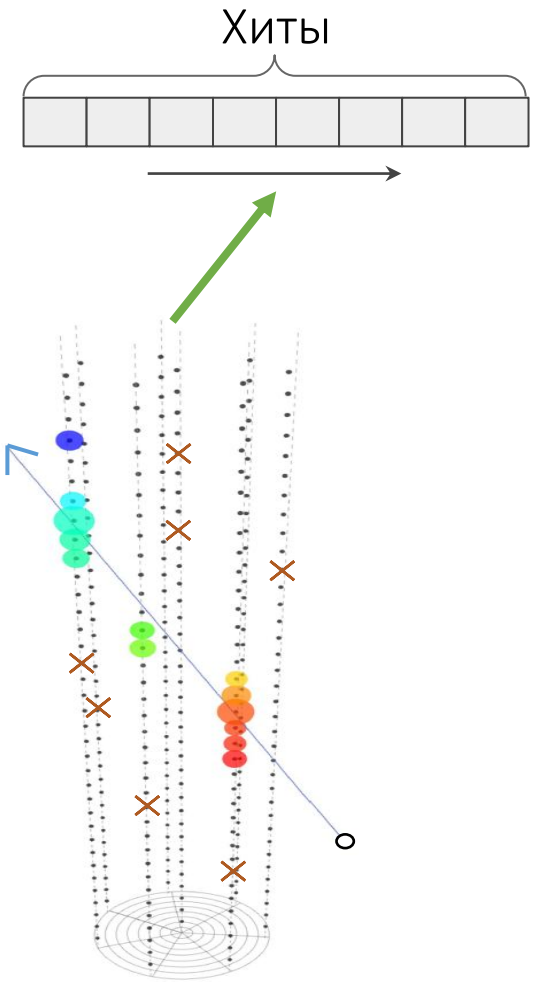
2. Схема применения нейросетей (NN)



2. Схема применения нейросетей (NN)



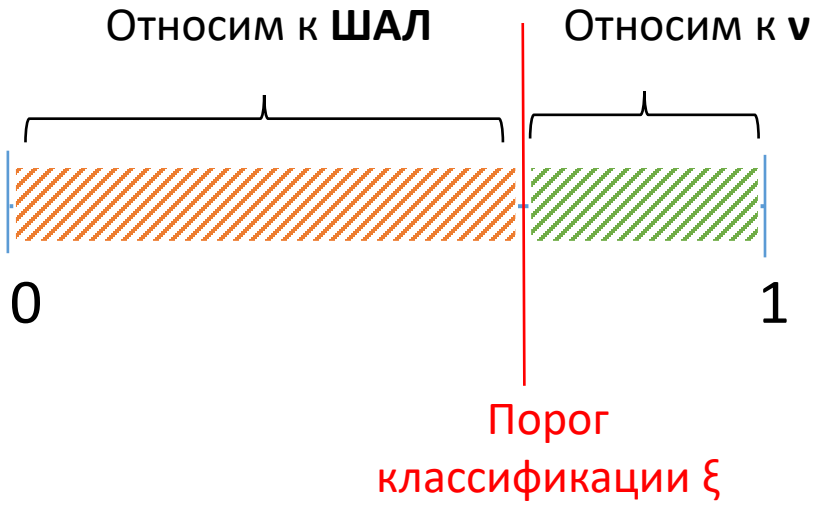
3. «Префильтр» событий



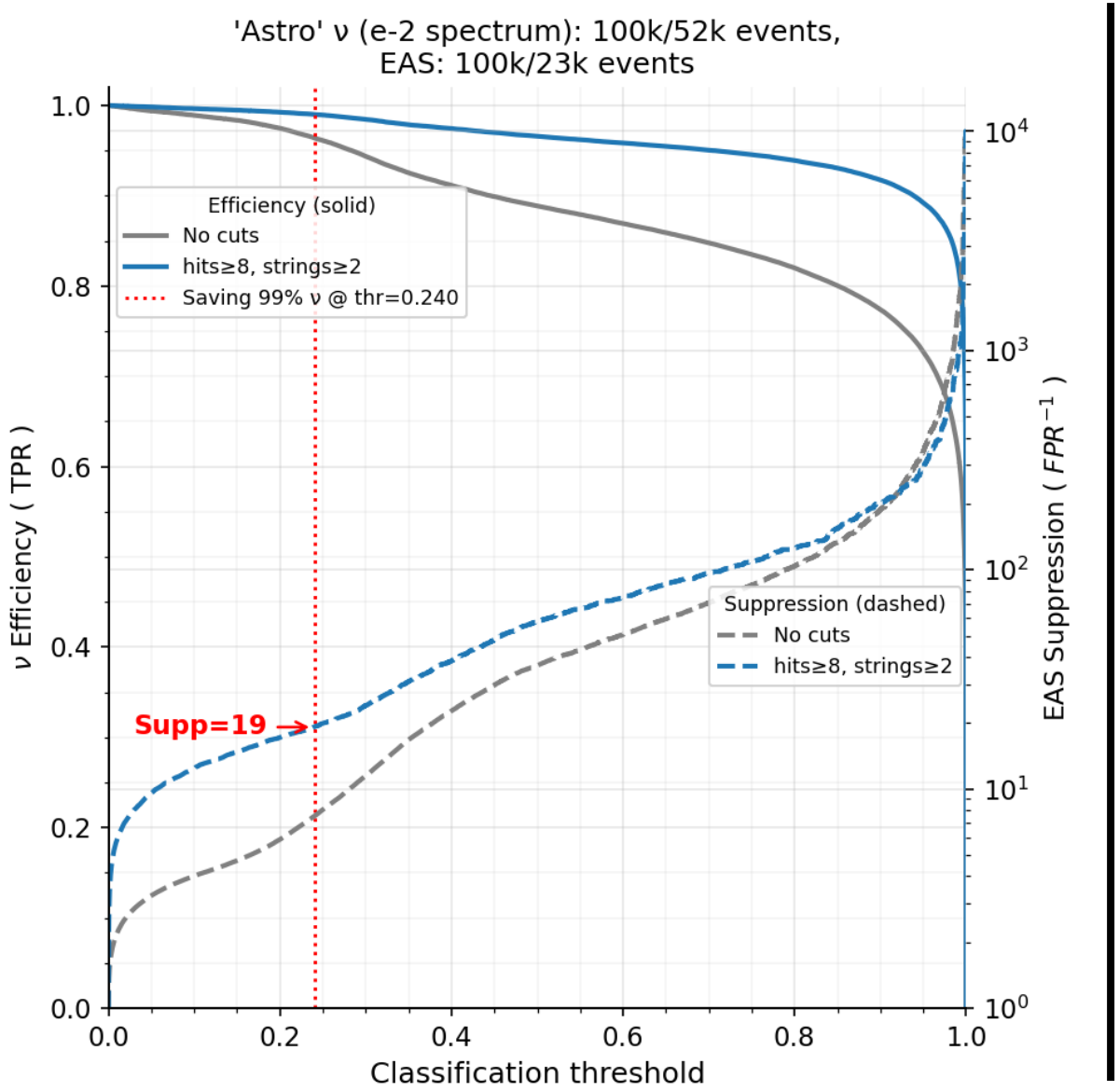
Transformer NN

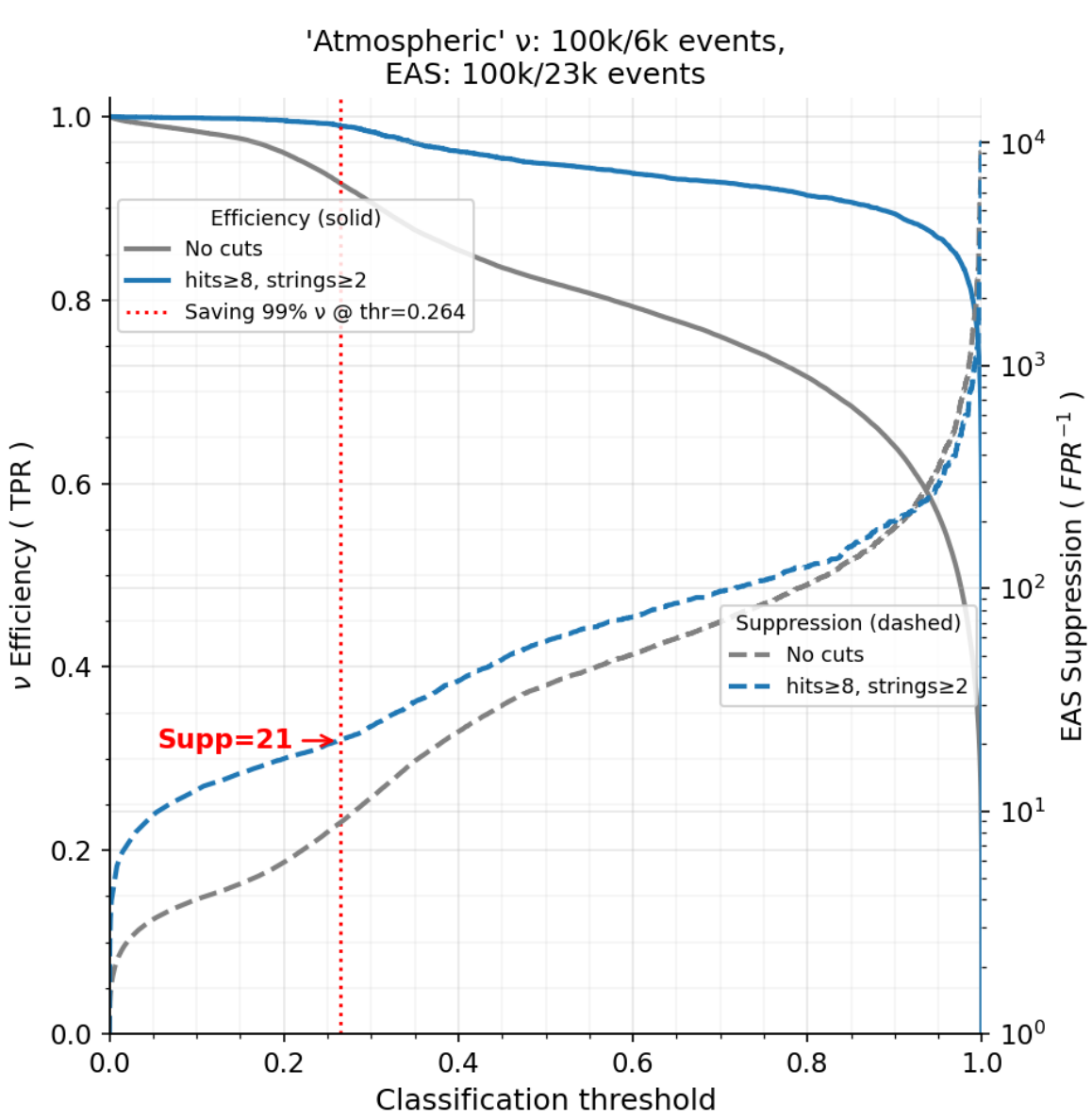
Предсказание p_ν

p_ν :



3. «Префильтр» событий: метрики на МК





Сохраняем 99% нейтрино, подавляем фон ШАЛ-ов в ~20 раз

3. Как адаптировать к реальным данным?

Модельная задача: распознавание цифр.

Train domain

Test domain

Цвет не важен,
но нейросеть этого не знает!

После обучения,
качество на цветных цифрах плохое!

3. Как адаптировать к реальным данным?

Модельная задача: распознавание цифр.

Train domain

Test domain

Цвет не важен,
но нейросеть этого не знает!

После обучения,
качество на цветных цифрах плохое!

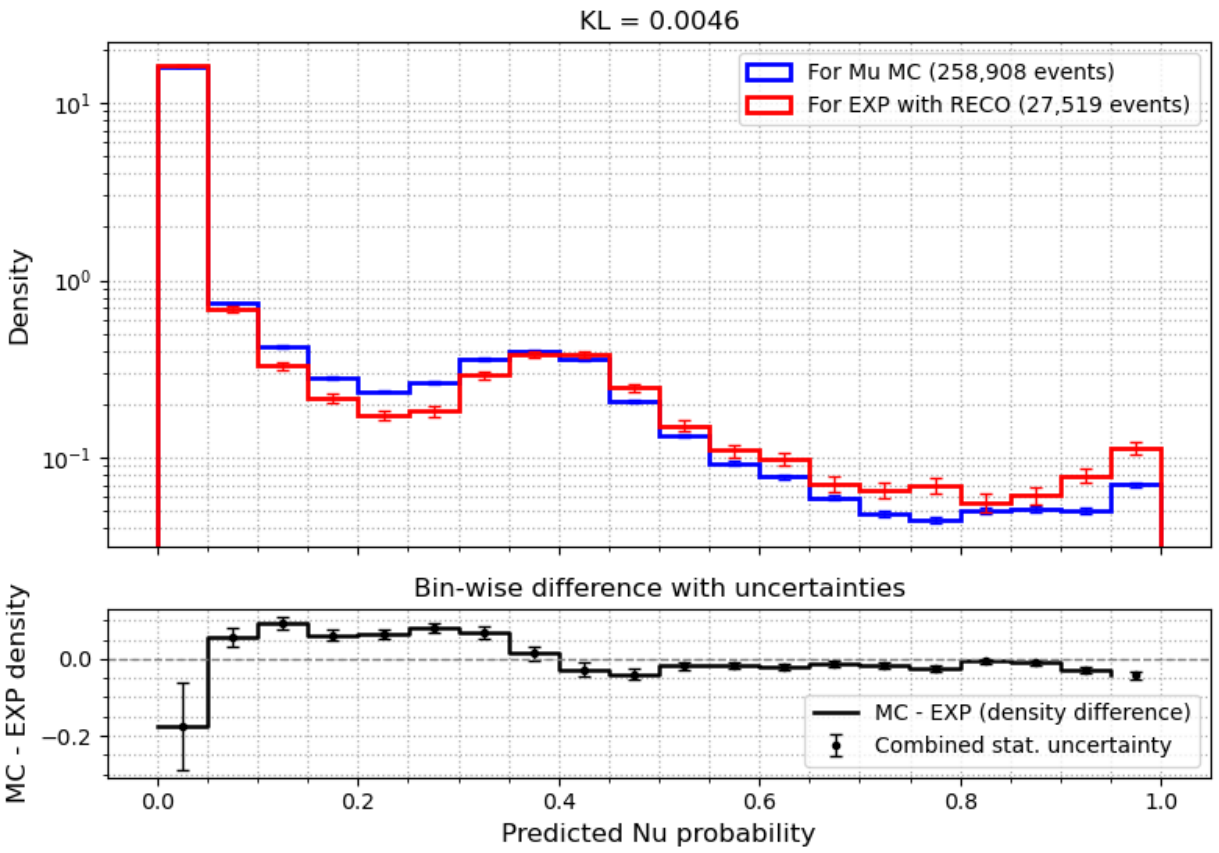
Данные для обучения и тестовые данные могут отличаться: разные домены
(у нас: МК vs эксперимент)

Но есть способ научить сеть извлекать домен-независимые признаки!
Доменная адаптация (ДА): arXiv:1409.7495

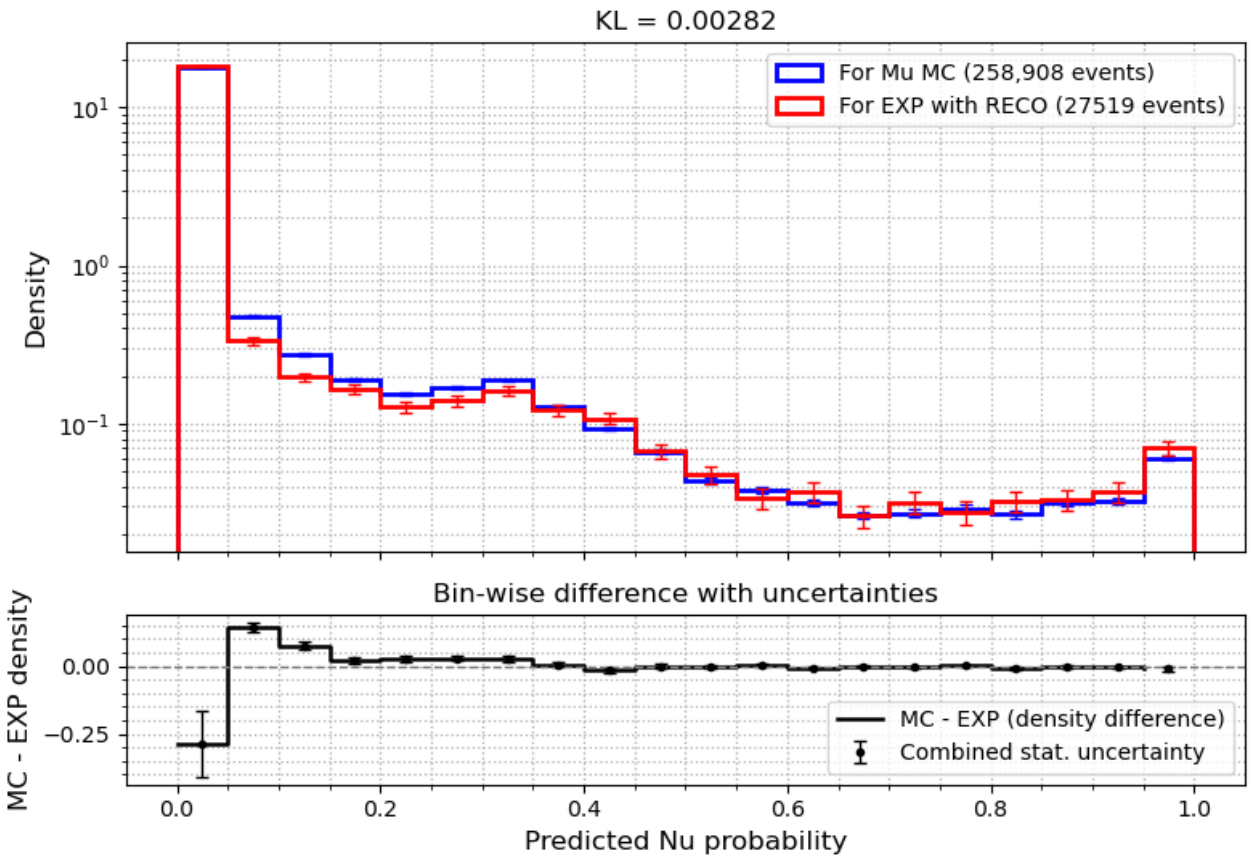
3. «Префильтр» событий: МК vs реальные данные

Сравнение предсказаний NN (каты h8s2)

Без Доменной Адаптации



С Доменной Адаптацией



С применением ДА распределение меньше зависит от домена данных!

KL дивергенция ↓

4. Другая задача: подавление шумовых хитов

Применения NN

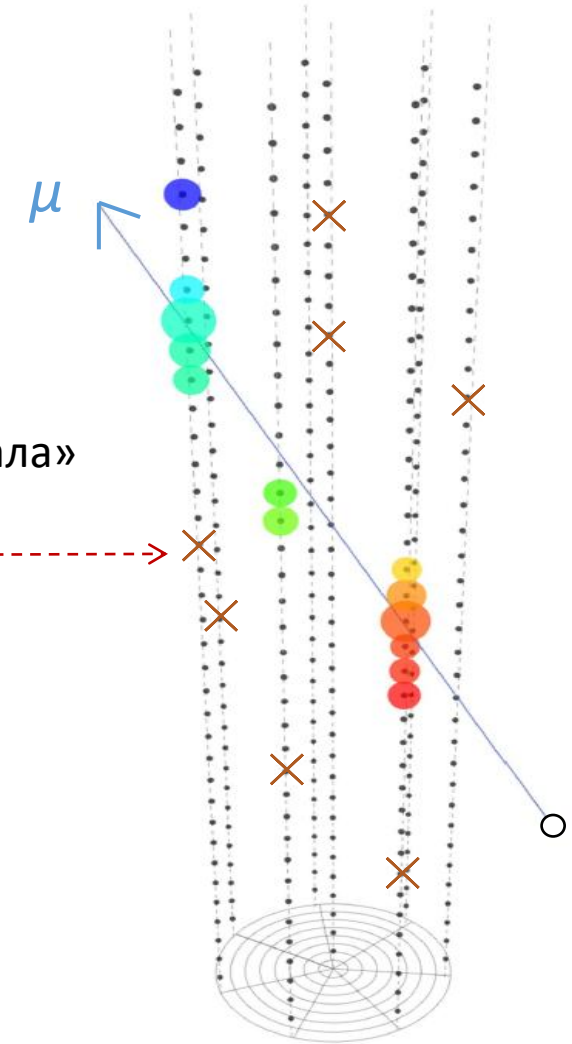
(1) «*Префильтр*» событий.
Цель: *грубо*, но *быстро* снизить фон

(2) *Подавление шумовых срабатываний*.
Цель: выделить только сигнальные хиты

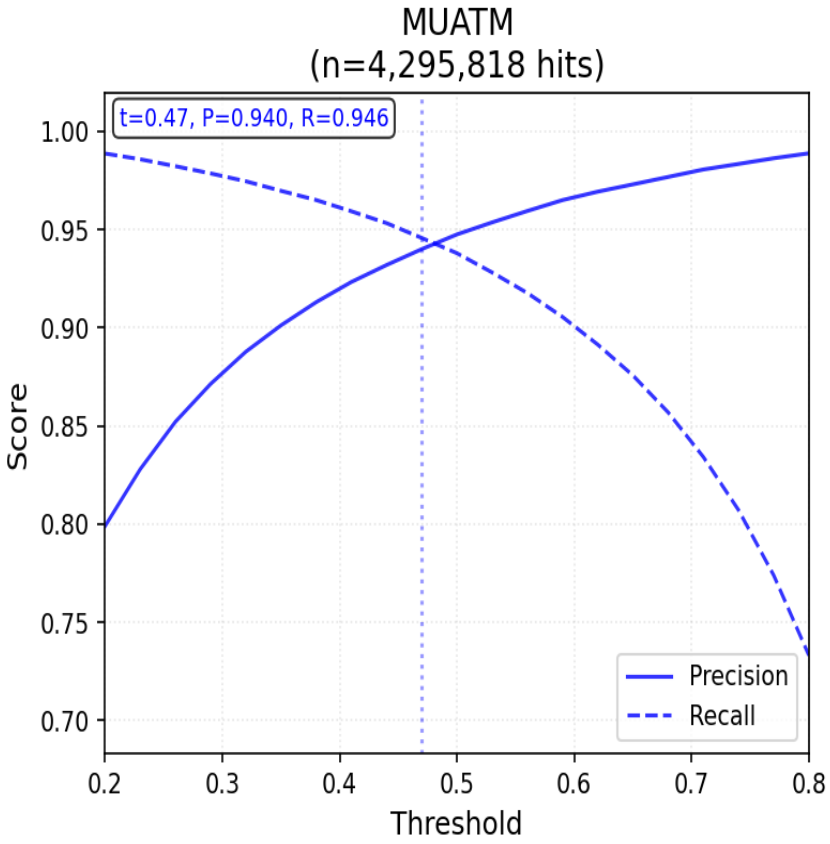
(3) *«Фильтр»* событий.
Цель: точный поиск событий от **ν**

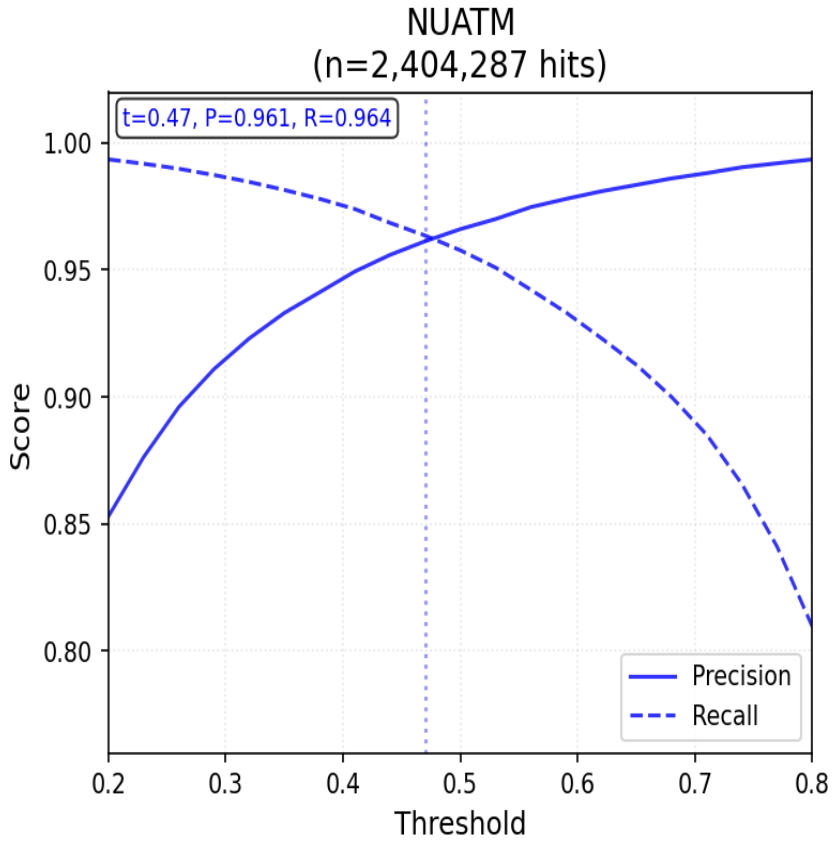
Для каждого хита:

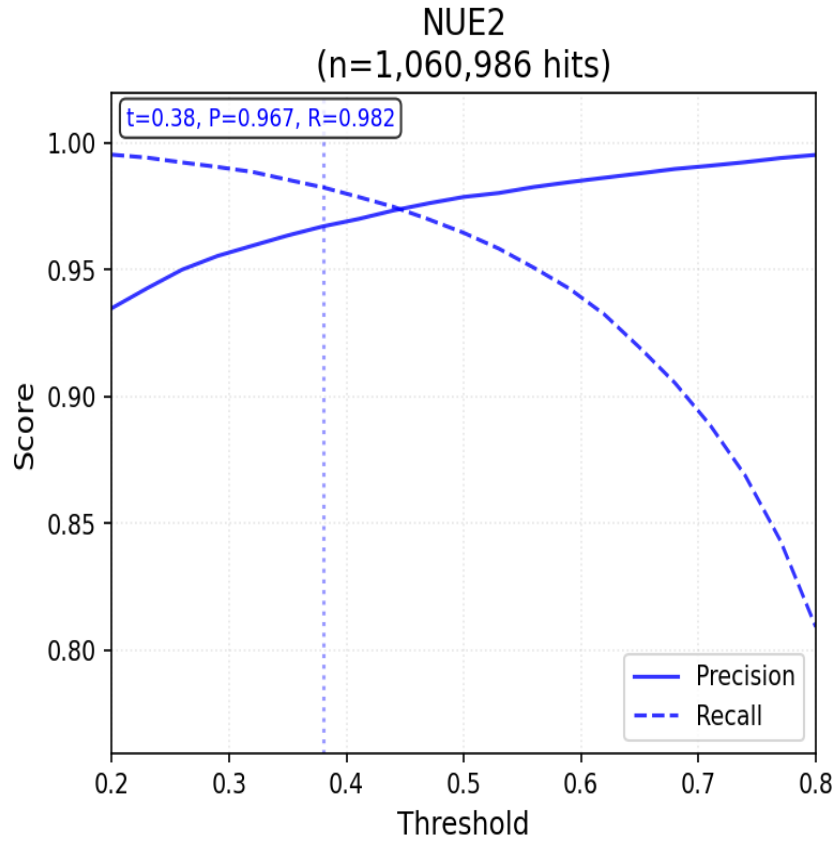
- оцениваем вероятность «сигнала»
- ставим кат



4. Подавление шумовых хитов: **метрики на МК**



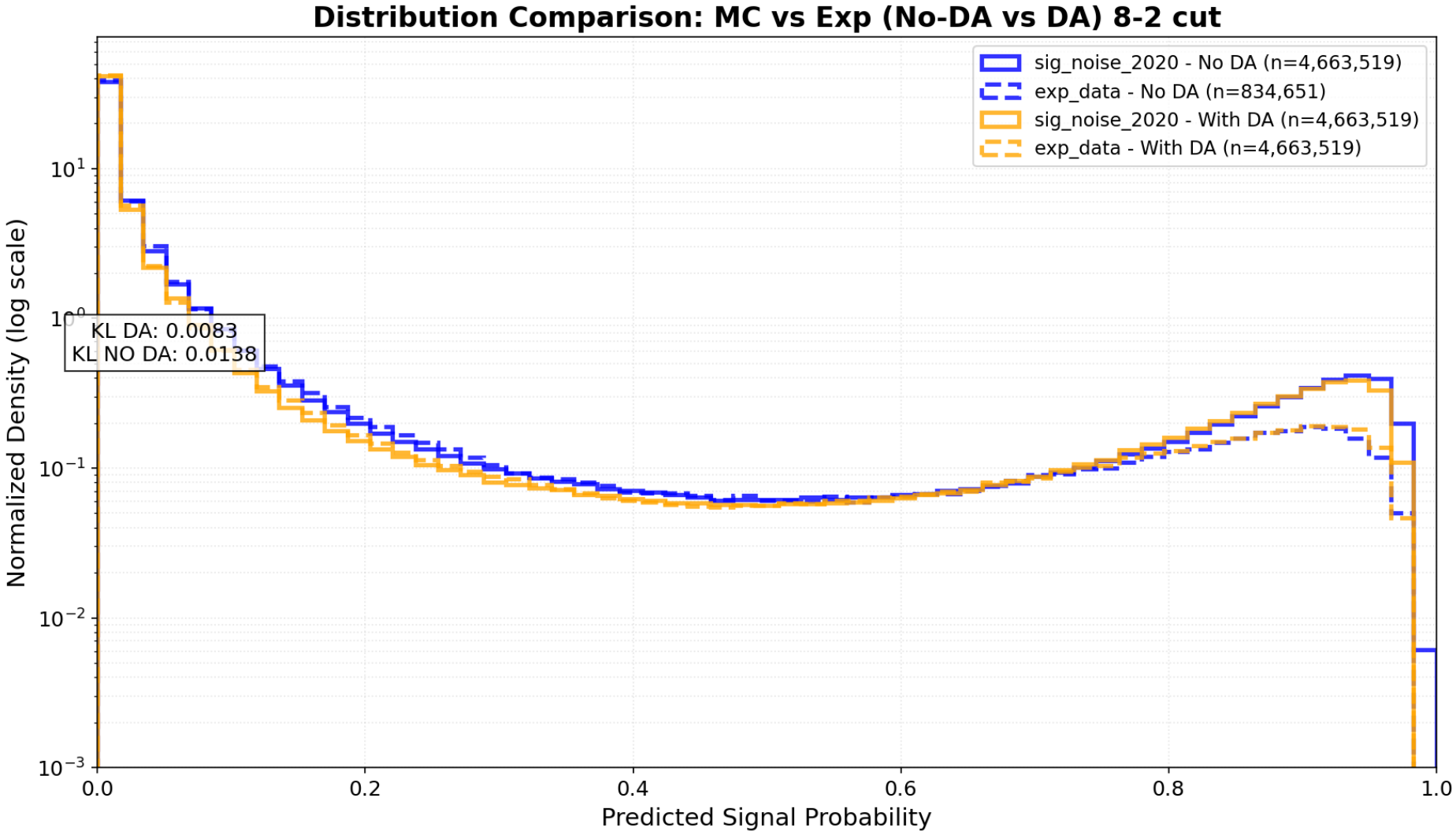




Recall (полнота, efficiency) – доля сохраняемых сигнальных хитов.
Достигаем Recall>95%!

Precision (точность, purity) – доля сигнальных хитов среди положительных предсказаний.
Достигаем Precision≈95%. При условии, что >80% хитов – это шум, подавление > 100 раз.

4. Подавление шумовых хитов: МК vs реальные данные



С применением ДА распределение (чуть-чуть) меньше зависит от домена данных.

KL дивергенция ↓

IV. Итоги

1. Есть прогресс в двух задачах (метрики по МК)

I. Префильтр событий
Сохраняем 99% h8s2 **ν** событий
Подавляем фон в 20 раз.

II. Подавление шумовых хитов
95% Полноты и Точности для всех хитов

2. Делаем шаги к реальному применению:

- Доменная Адаптация делает предсказания моделей устойчивее при переходе от МК к домену реальных (EXP) данных

3. Дальнейшие планы:

- Проверка общего пайплайна:
  - Объединяем модели ***I.*** и ***II.,*** чтобы *ставить каты на EXP данные*
  - → сравниваем предсказания на MC and EXP
- Решаем задачу точного отбора событий-кандидатов **ν**

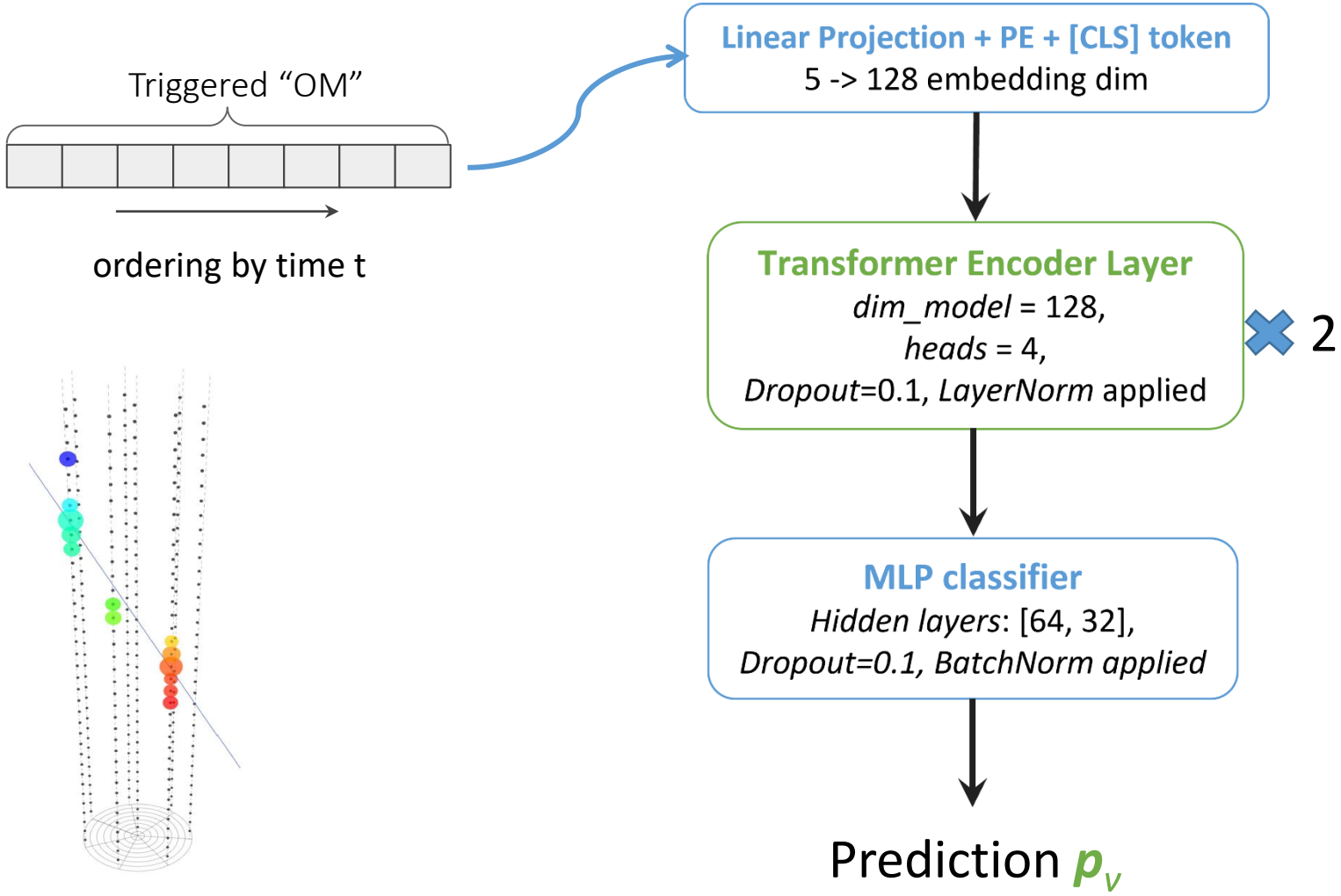
Спасибо!
Ваши вопросы?

Контакты:
matseiko.av@phystech.su
t.me/AlbertMac280

Код:
github.com/ml-inr/Baikal-ML

Backup

Fast pre-filter: The network



Training details:

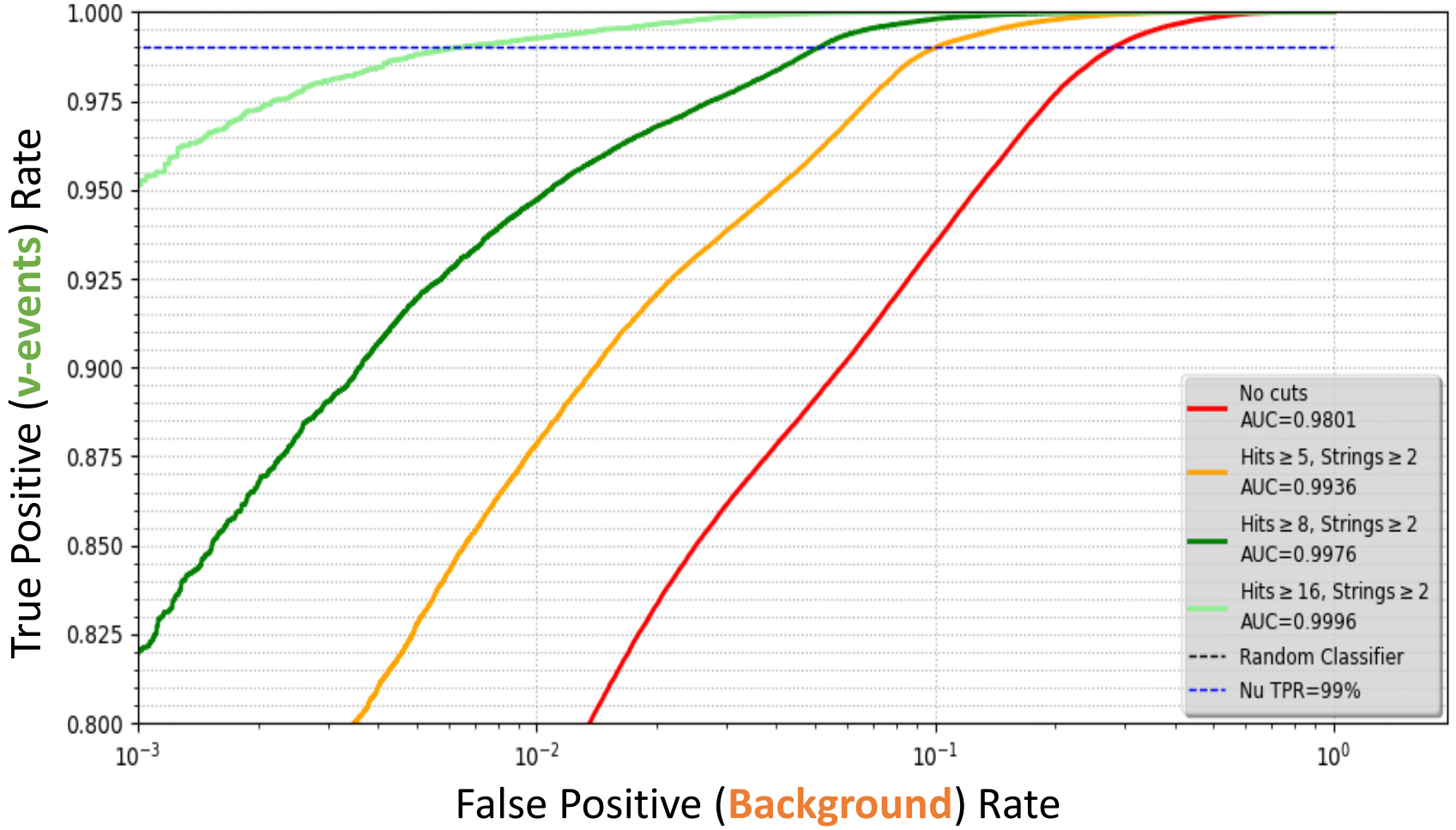
- AdamW optimizer
- BCE loss
- LR=2e-3; exp decay with 0.01 weight
- Early stopping by validation ROC AUC, patience 15 epochs

3. «Префильтр» событий
ROC Curves

2020 year MC Data:
- 500,000 **muatm** events
- 500,000 **ν** events
- **ν-atm** and **νe-2** samples in equal

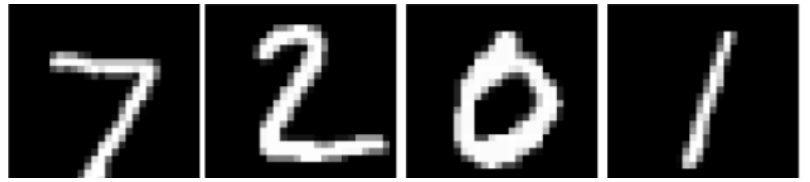
Results for h8s2 events:
- saving 99% **ν** events
- suppressing background by factor 20

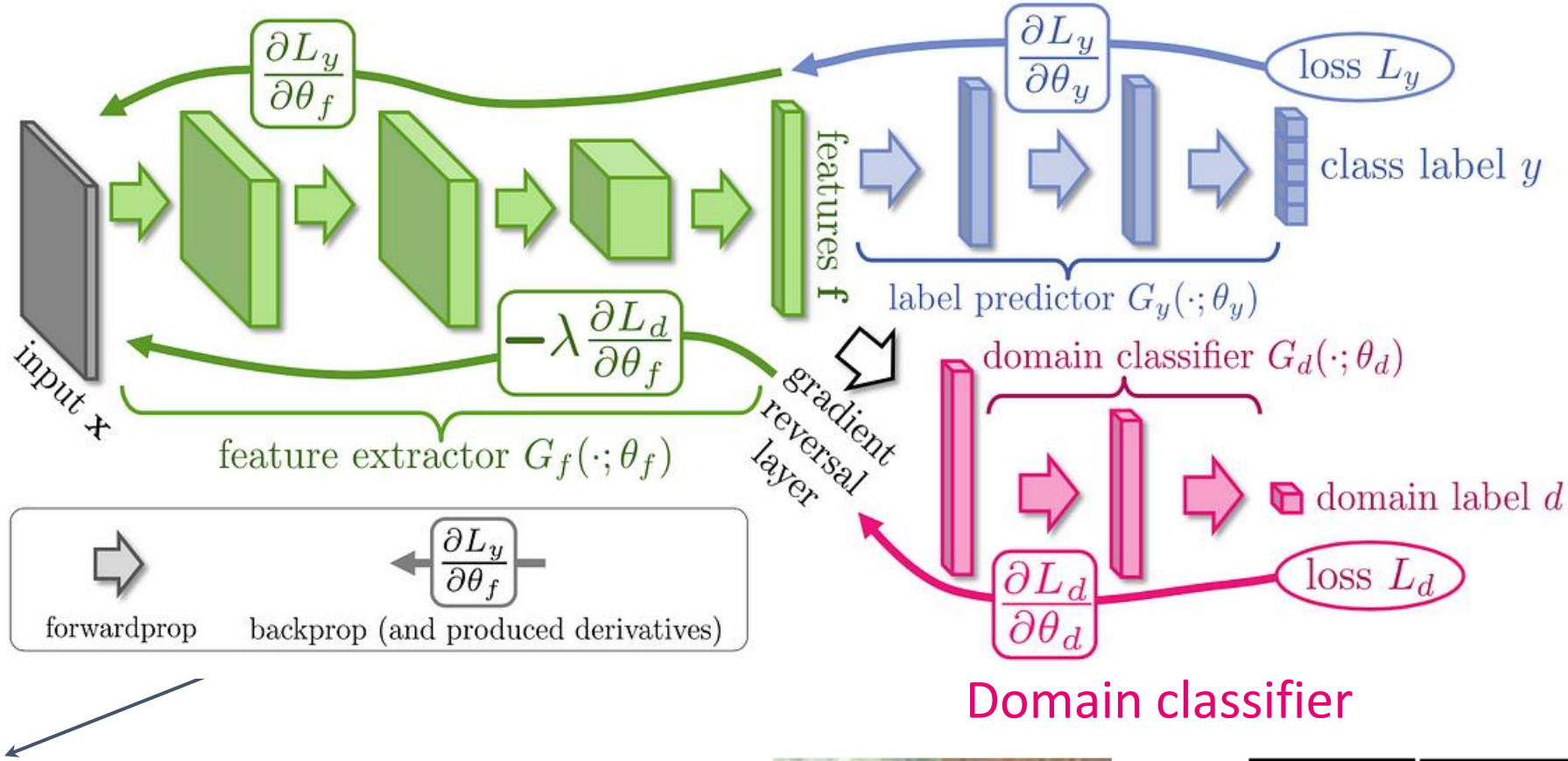
(threshold 0.241)



Domain adaptation

Digit recognition





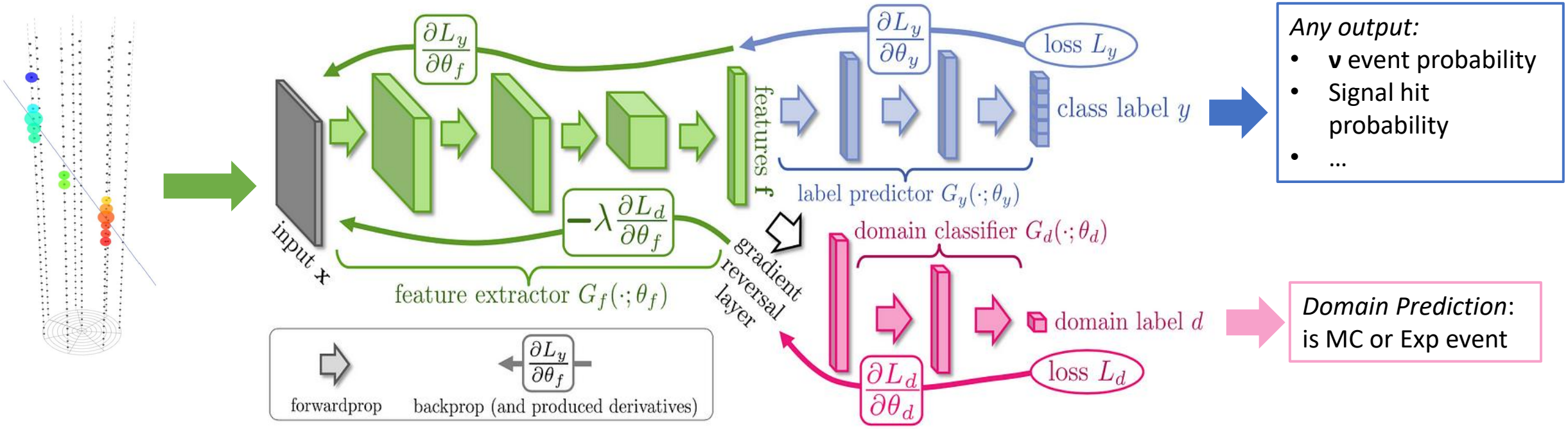
Domain classifier

Gradient reversal:
Exclude domain-specific features

vs

2.4 Domain Adaptation Technique

DA scheme:
domain-adversarial training with a gradient reversal layer



Any output:
- **ν** event probability
- Signal hit probability
- …

Domain Prediction:
is MC or Exp event

Source arXiv:1505.07818

2.2 Experiment data overview

“Raw” Events

- “Good” runs are taken equally distributed over the 2020 year
- Only the first 25,000 events were taken from each run!
- No cuts!
- About 650,000 events in total

Season	Cluster Number	Run Numbers
2020	1	27, 116, 188
2020	2	15, 149, 206, 357
2020	3	12, 115, 263, 417
2020	4	37, 116, 117
2020	5	35, 116, 270, 377
2020	6	32, 134, 325, 413
2020	7	41, 120, 202, 406

Reconstructed Sample

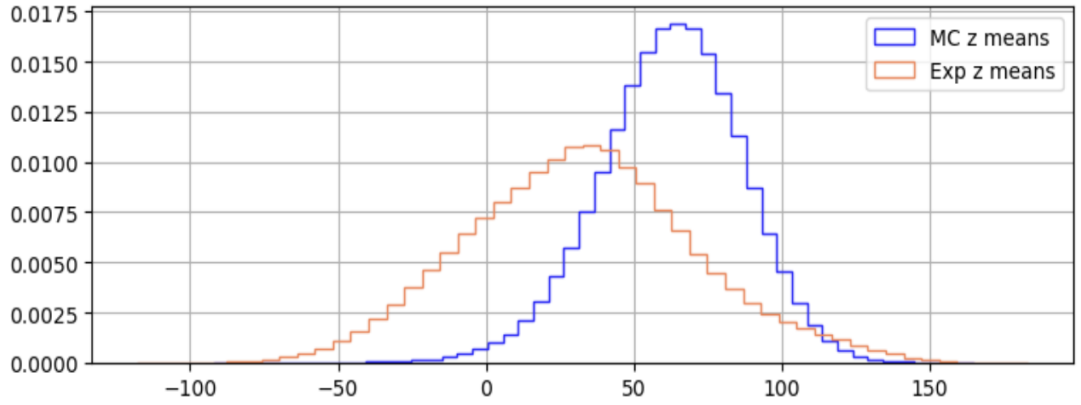
- Shared by Grigory Safronov
- Applied cuts:
  - min signal 8 hits
  - min 2 signal strings
- 27,530 events

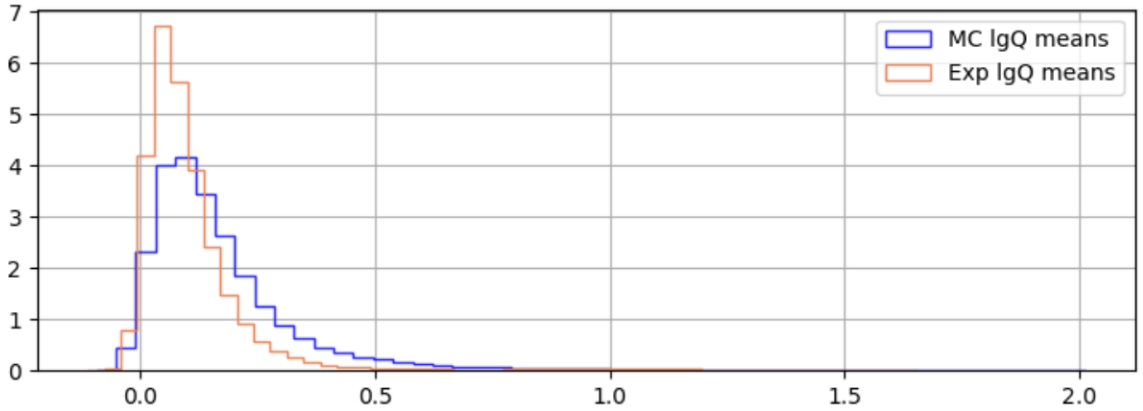
Season	Cluster Number	Run Numbers
2020	3	120 -- 129

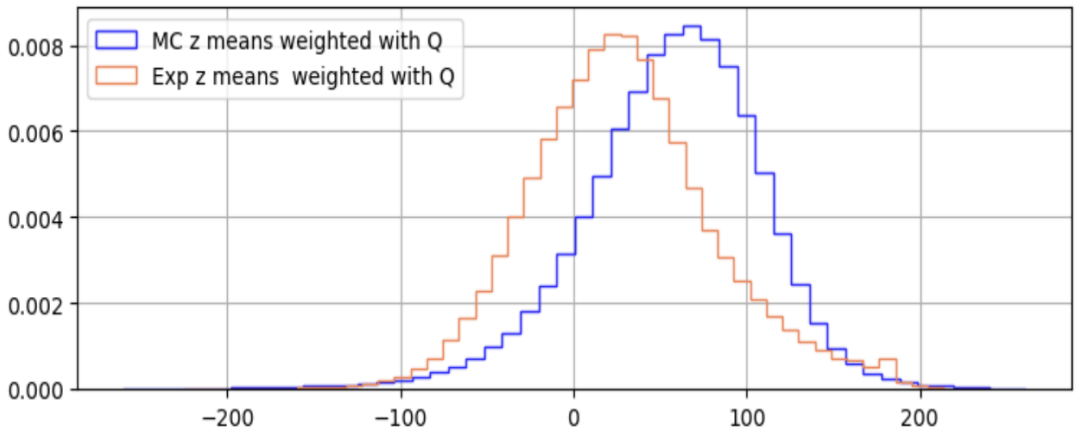
2.3 Exp vs muatm MC Data

Observed discrepancies for Z and Q

Possible reason: many events with noise hits only in Exp Data







How statistics are obtained:

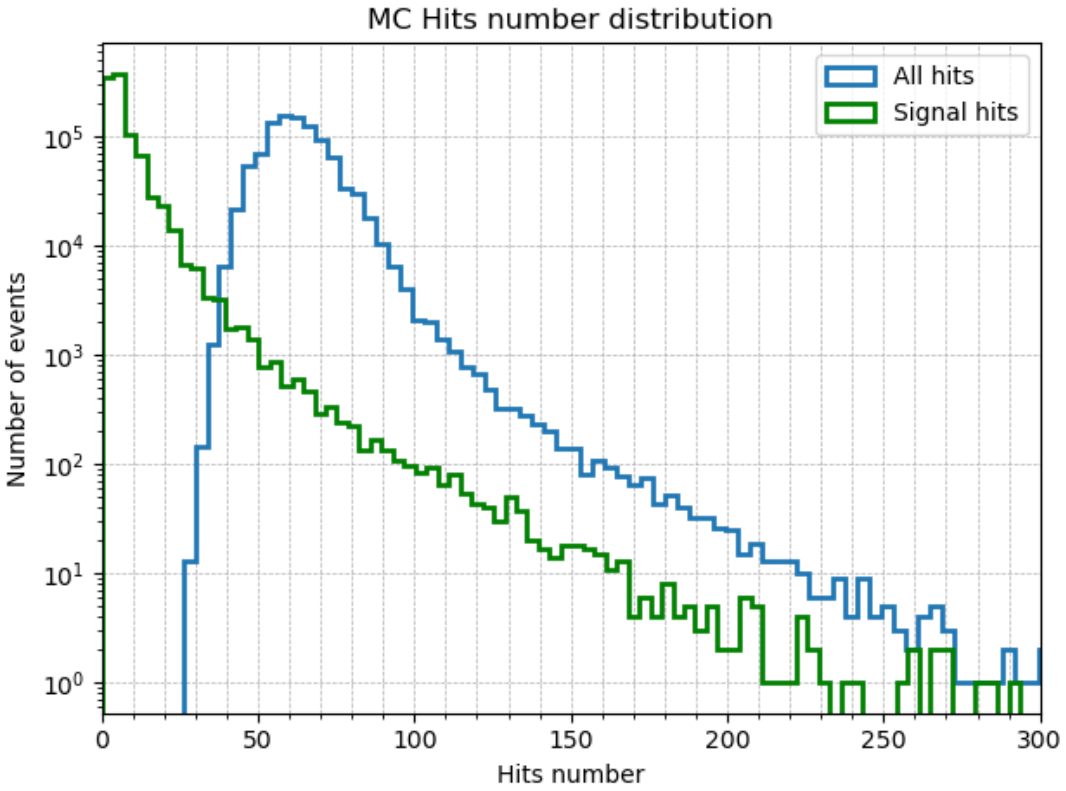
- Coordinate system is shifted to the cluster's center
- *Means* of Z and Q were calculated *for each event*

Fast pre-filter: dataset

- Using **Monte-Carlo simulation**[1]
  - EAS evolution and propagation of particles in water.
  - Random noise hits
  - Almost fixed geometry: no strings floating

- Track events only:
  1) Muons from EAS
  2) ν_μ (neutrinos of muon flavour)

- **Target feature** — type of particle
  Labels: 0 — EAS, 1 — neutrino



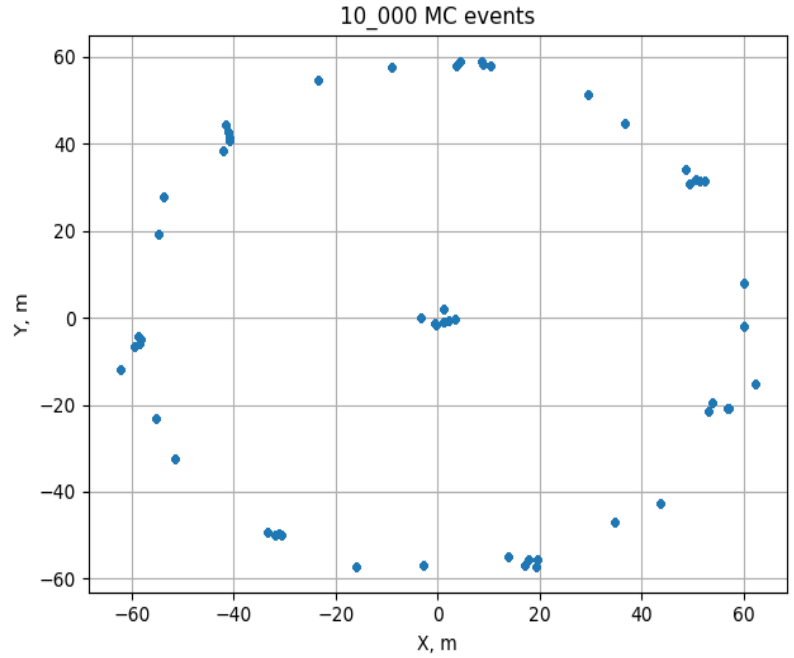
	Train	Validation	Test
EAS	$2.4*10^5$	$6*10^4$	$\approx 5*10^5$
Neutrino	$2.4*10^5$	$6*10^4$	$\approx 5*10^5$

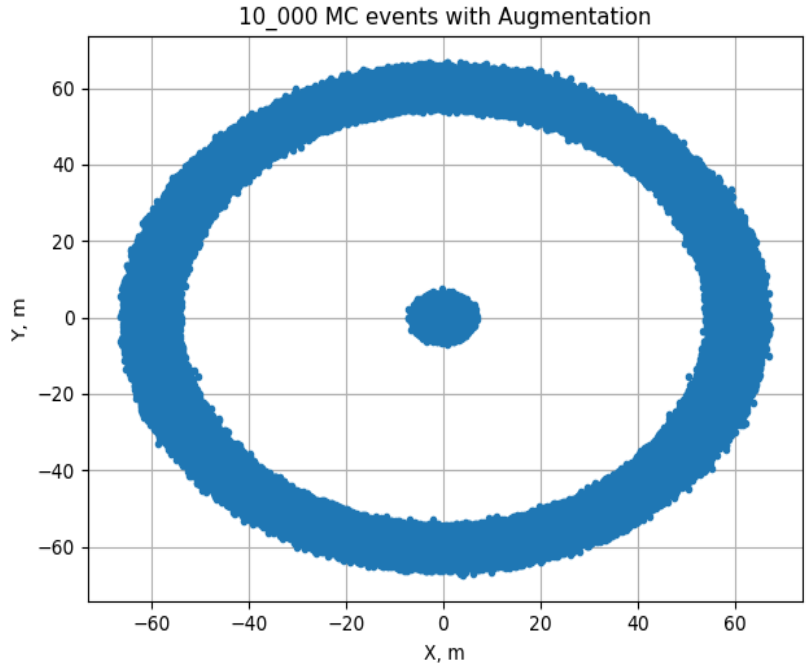
[1] arxiv.org/abs/2106.06288

Fast pre-filter: dataset

Data augmentations:

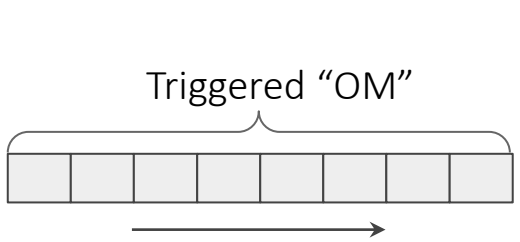
- Random gaussian noise to all inputs.
  STDs:
  - 1 m for **x** and **y**
  - 2m for **z**
  - 5 ns for **t**
  - 0.05 pe for **Q**
- Random rotations in x-y plane

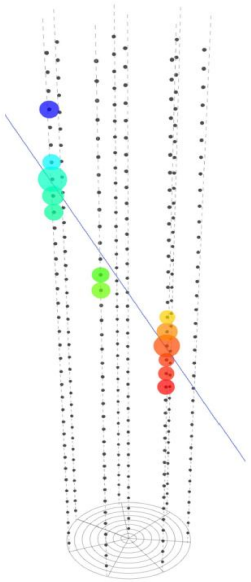


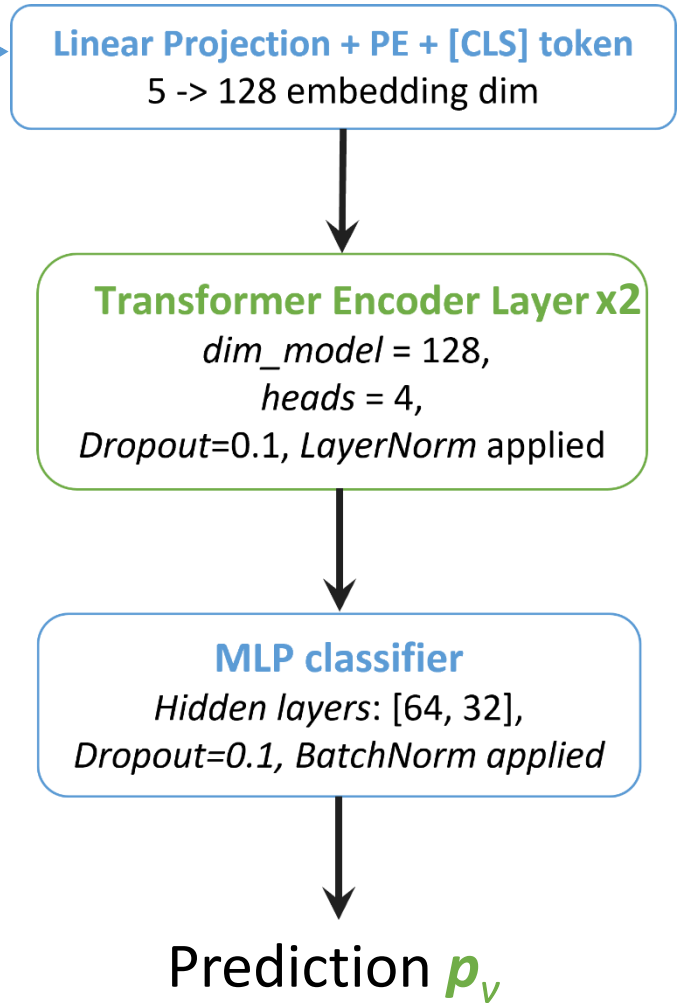


[1] arxiv.org/abs/2106.06288

Fast pre-filter: The network





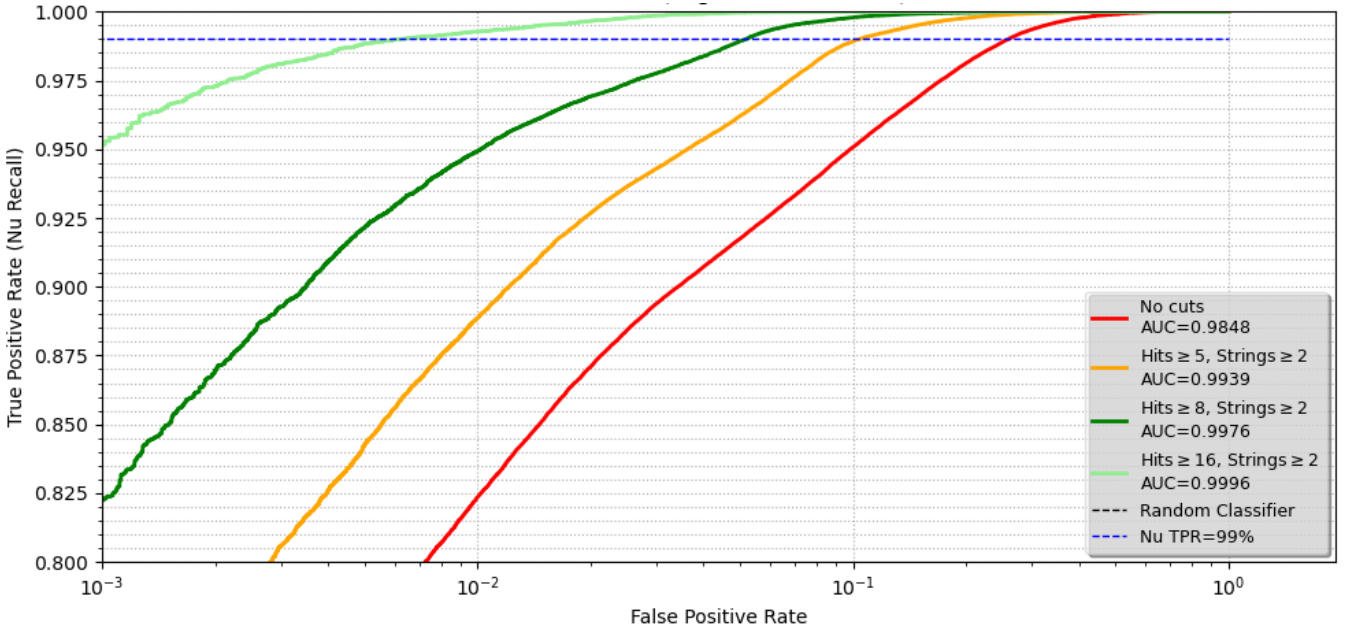


Training details:

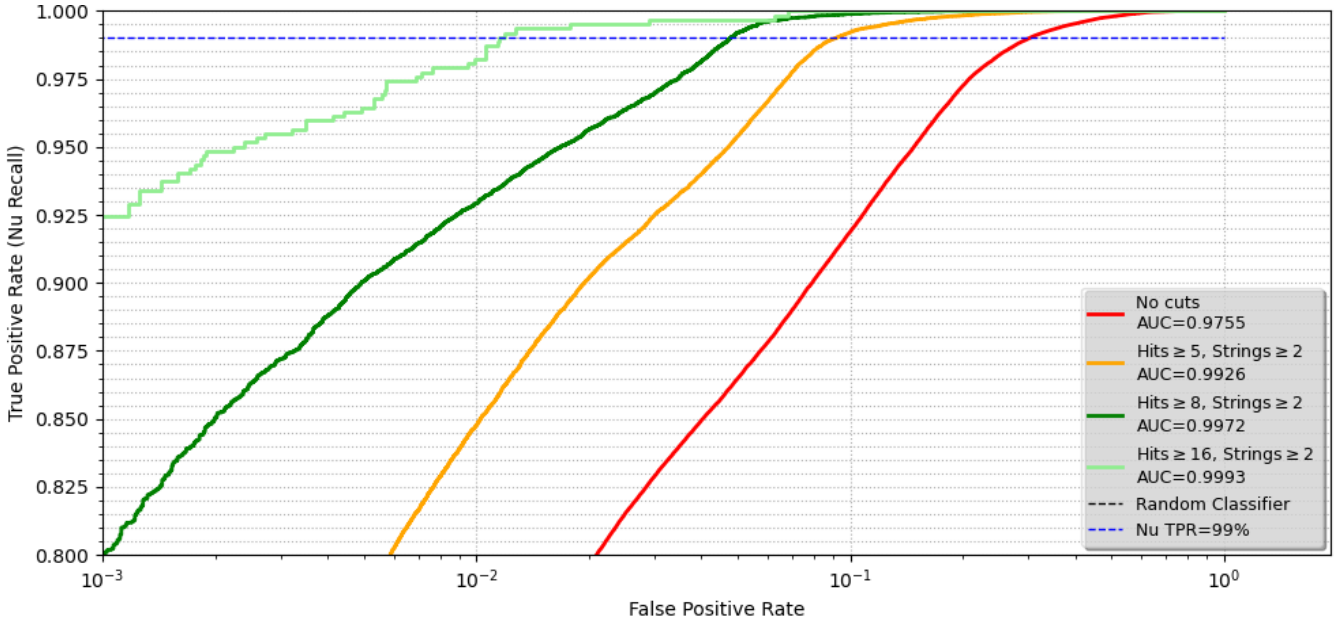
- AdamW optimizer
- BCE loss
- LR=2e-3; exp decay with 0.01 weight
- Early stopping by validation ROC AUC, patience 15 epochs

Fast pre-filter: more metrics

Nu Astro

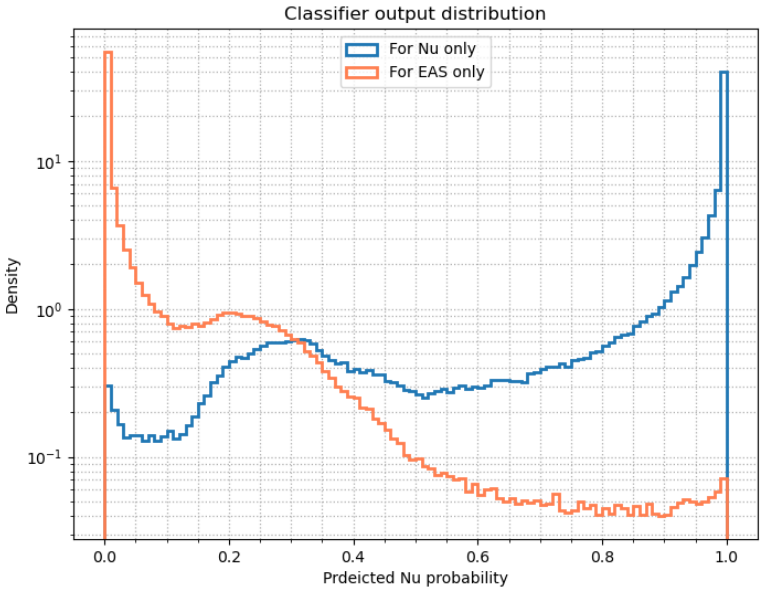


NuAtm

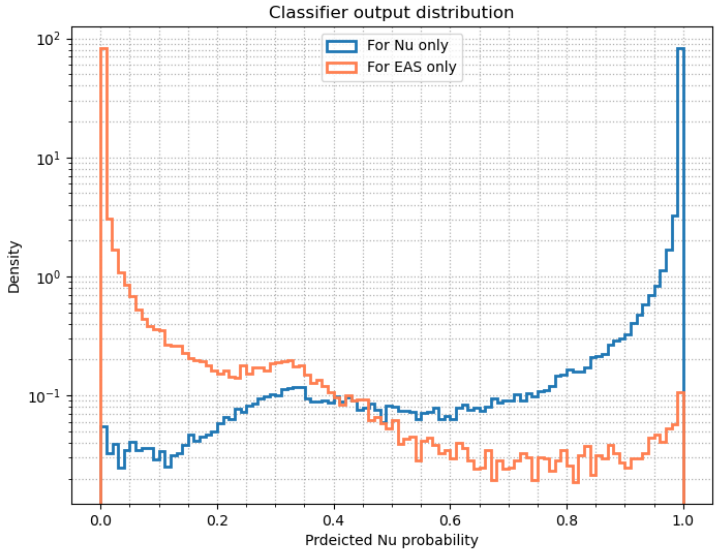


More “pre-filter” graphs

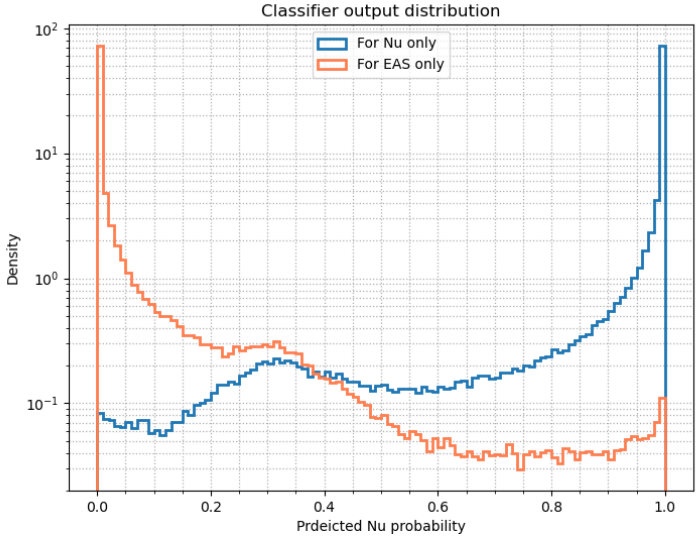
h0s0



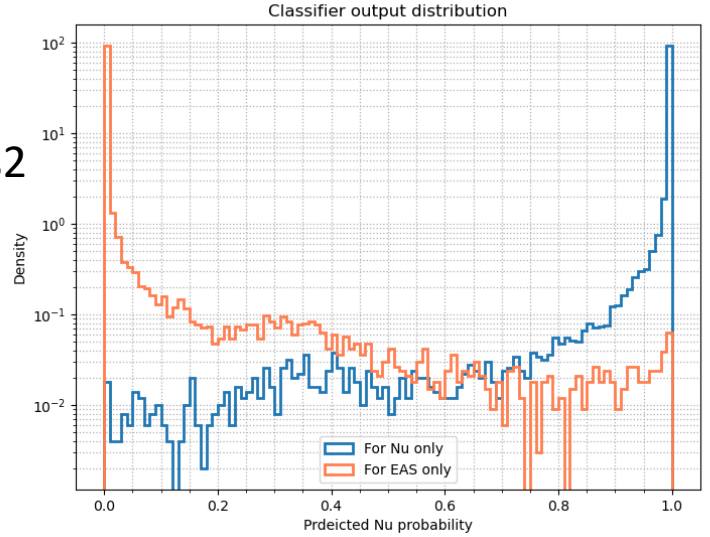
h8s2



h5s2



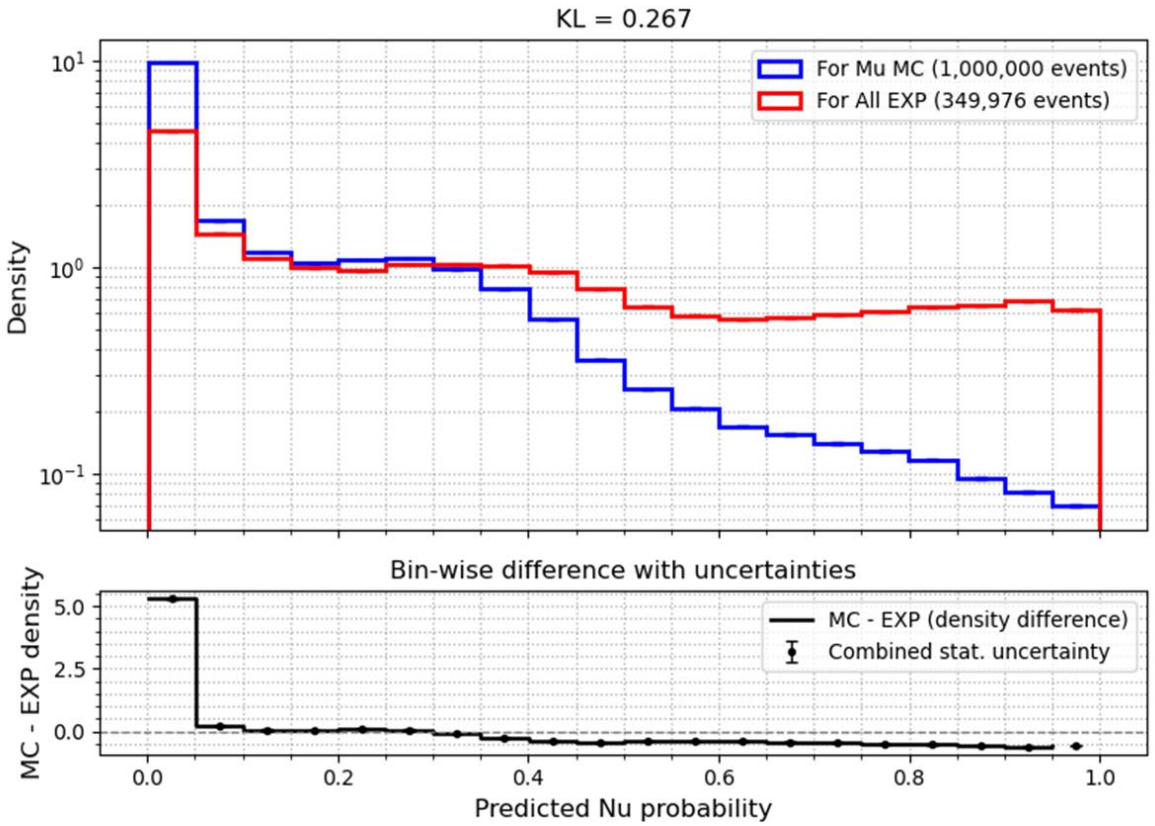
h16s2



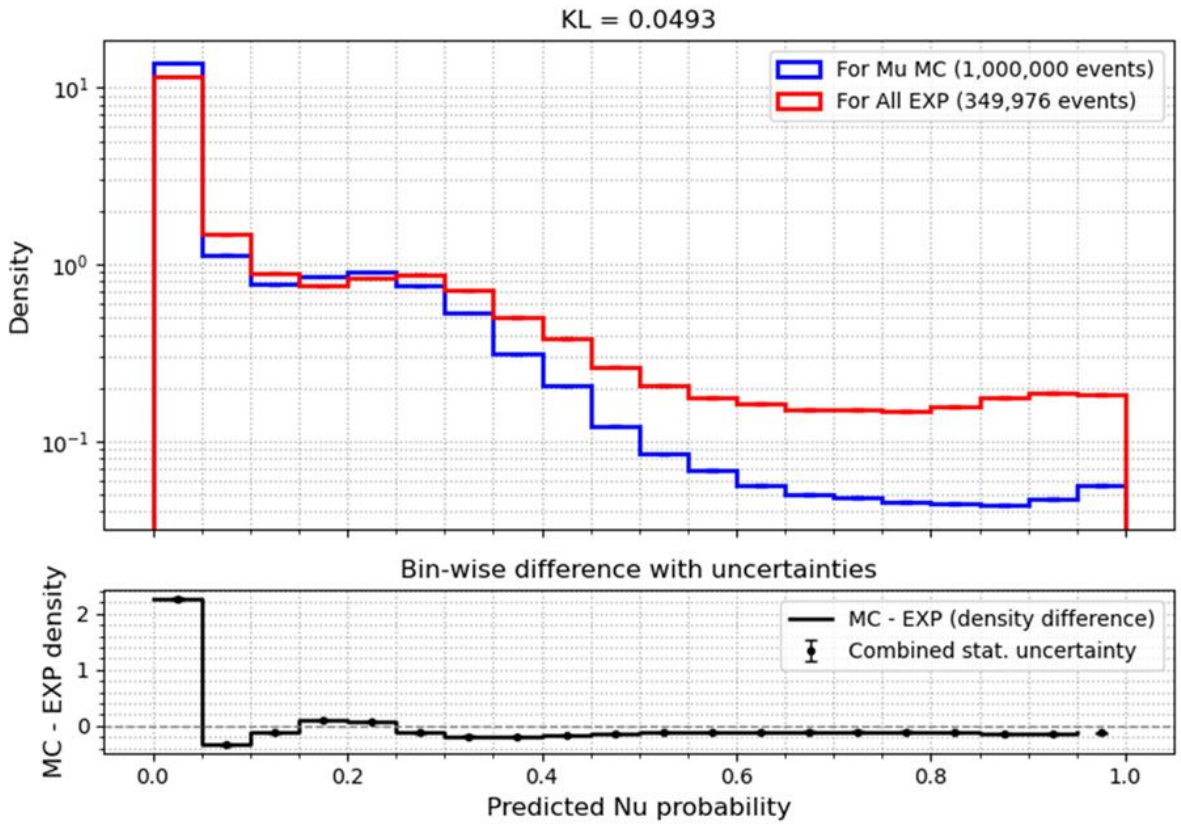
2.4 DA for Background Suppression

NN output distribution (No cuts)

No Domain Adaptation

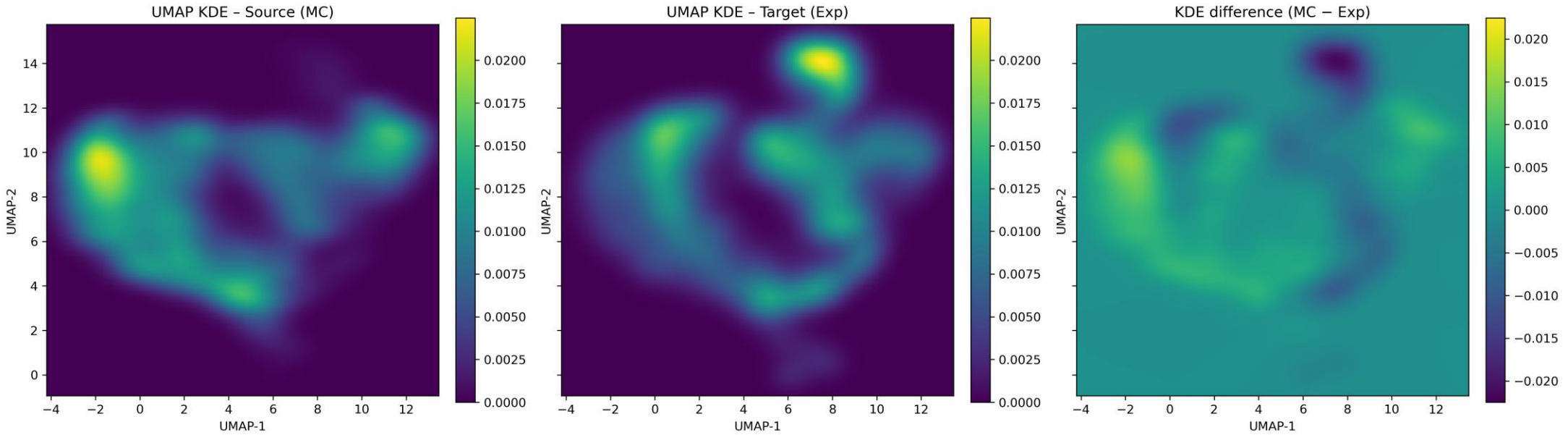


With Domain Adaptation

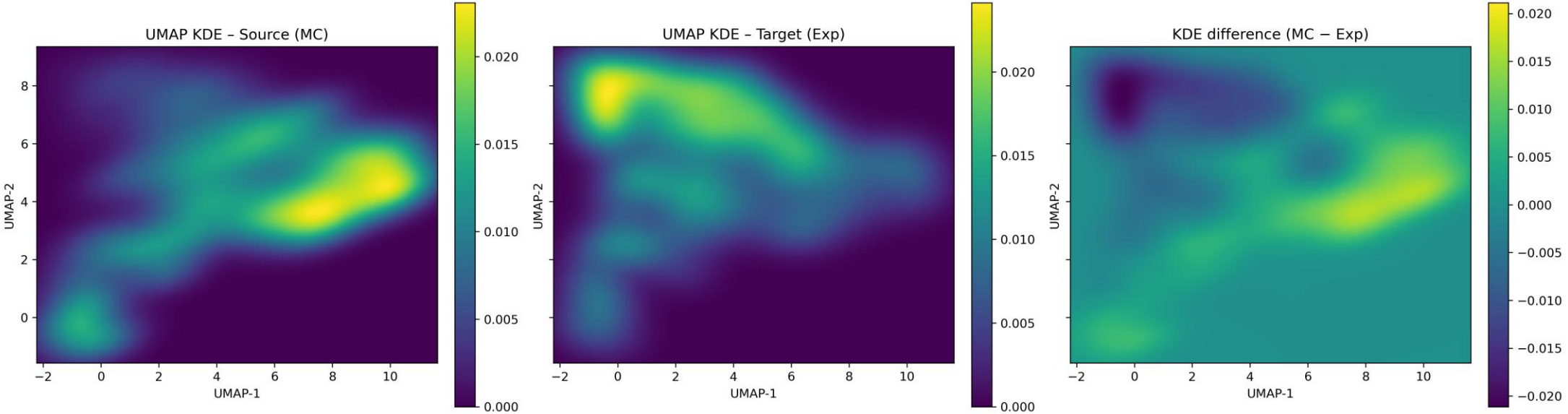


Fast pre-filter: UMAP KDE Diff

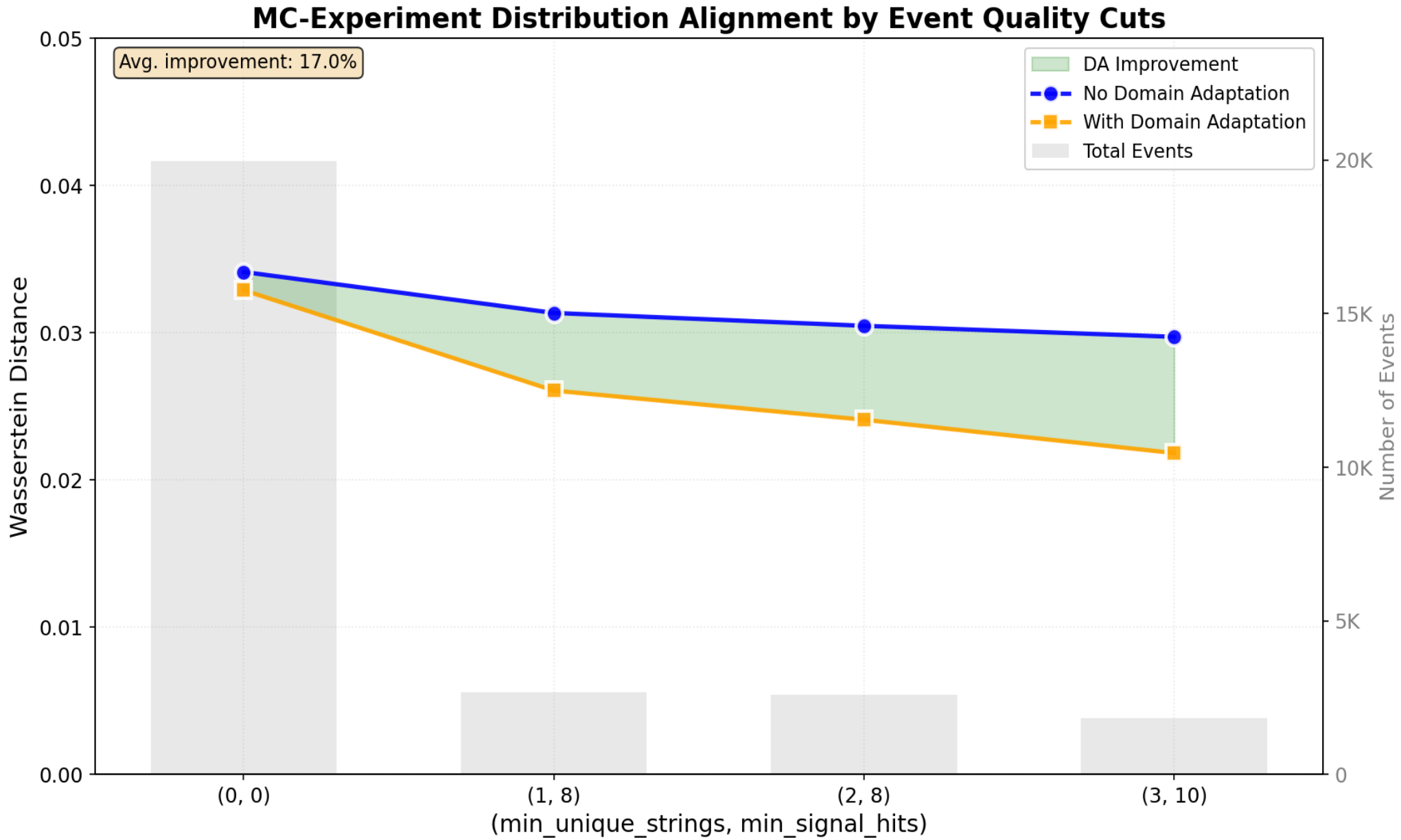
DA



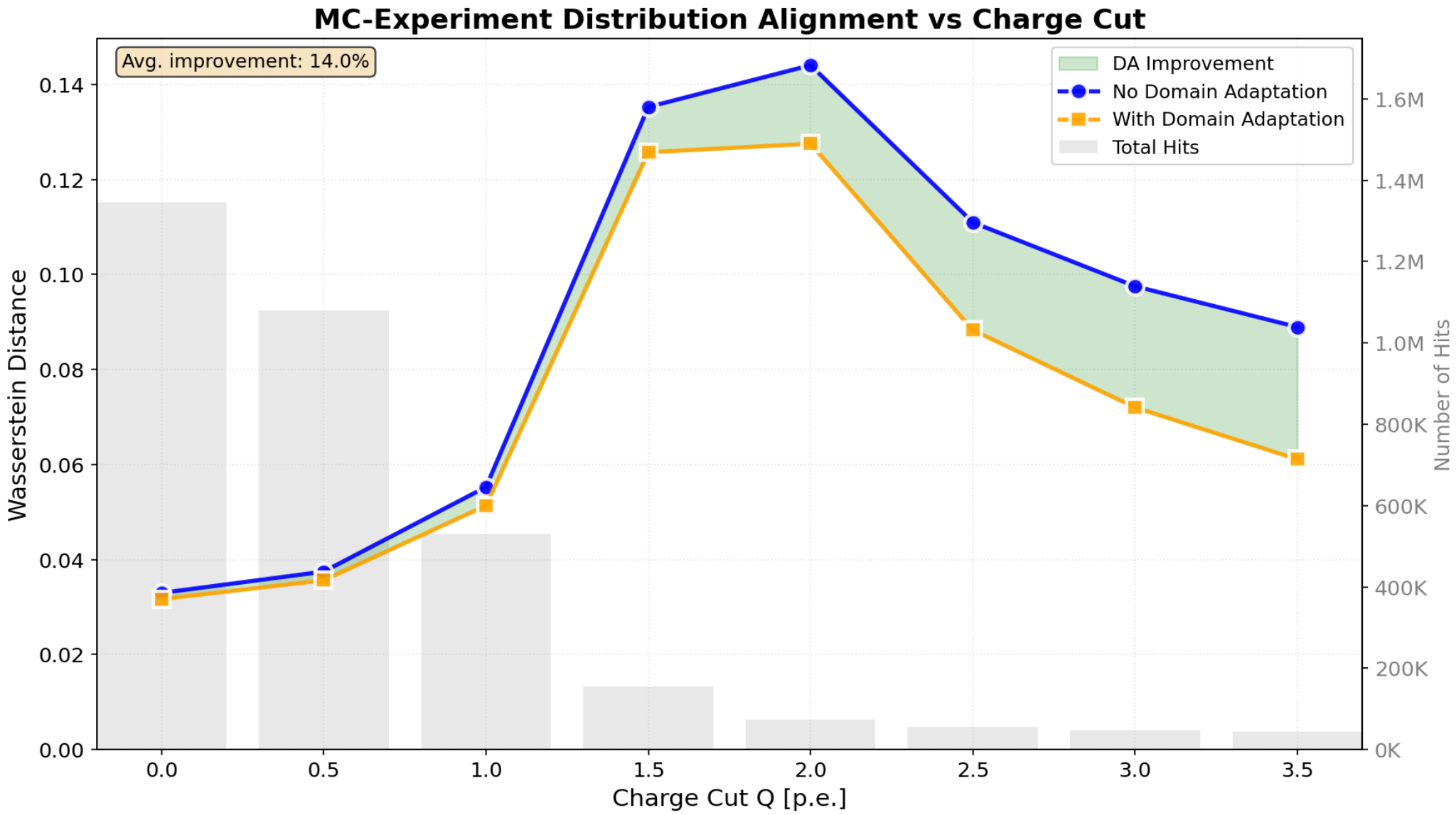
No DA



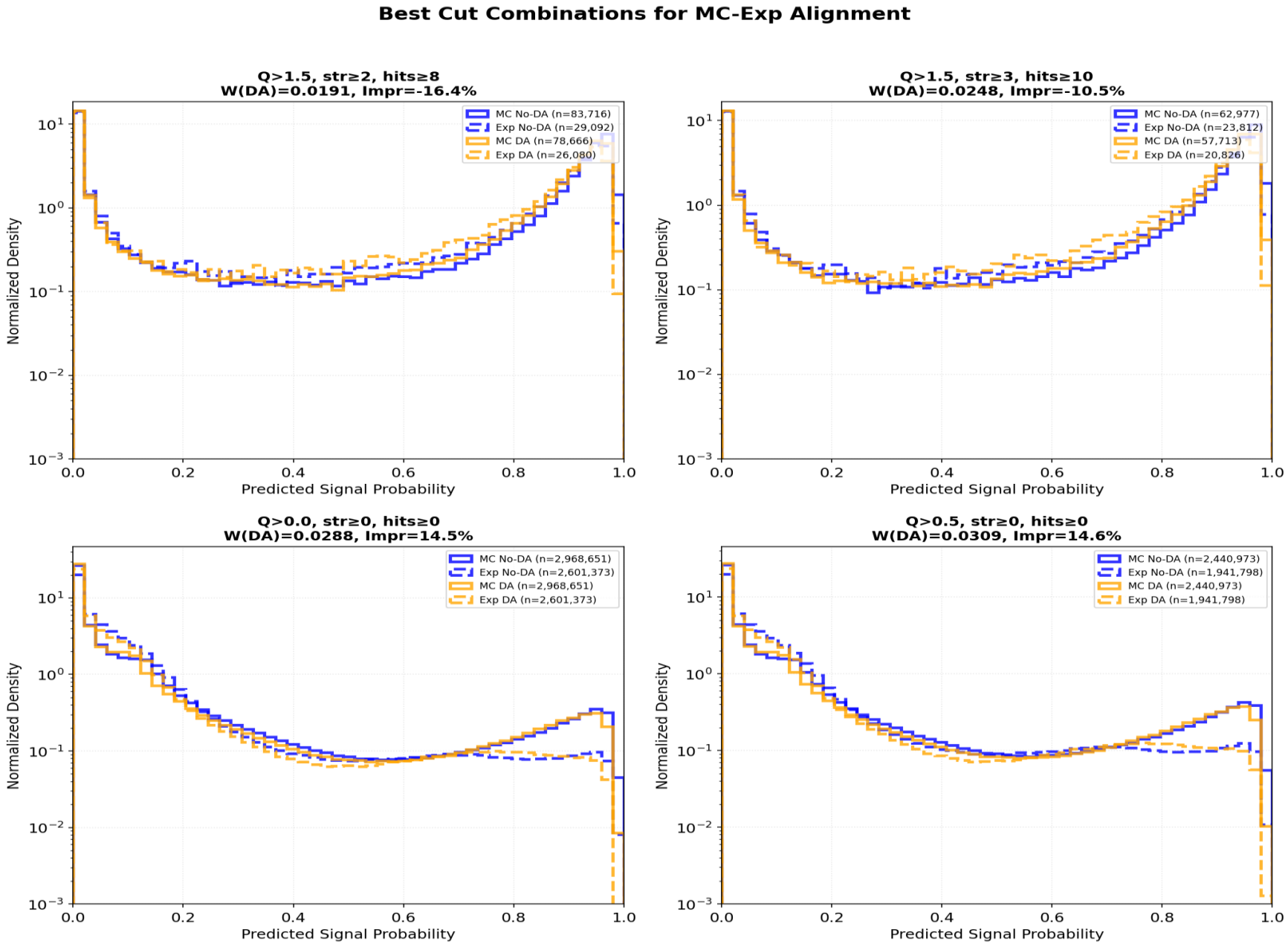
3.2 Real Data Results: DA for Noise Hits Suppression



3.2 Real Data Results: DA for Noise Hits Suppression



3.2 Real Data Results: DA for Noise Hits Suppression



Noise Hits Suppression: UMAPs