



Институт искусственного интеллекта и
цифровых наук
Факультет компьютерных наук

Лаборатория методов анализа
больших данных (LAMBDA)

Дубна 2024

Участие группы ВШЭ в экспериментах на БАК

Денис Деркач



Структура группы

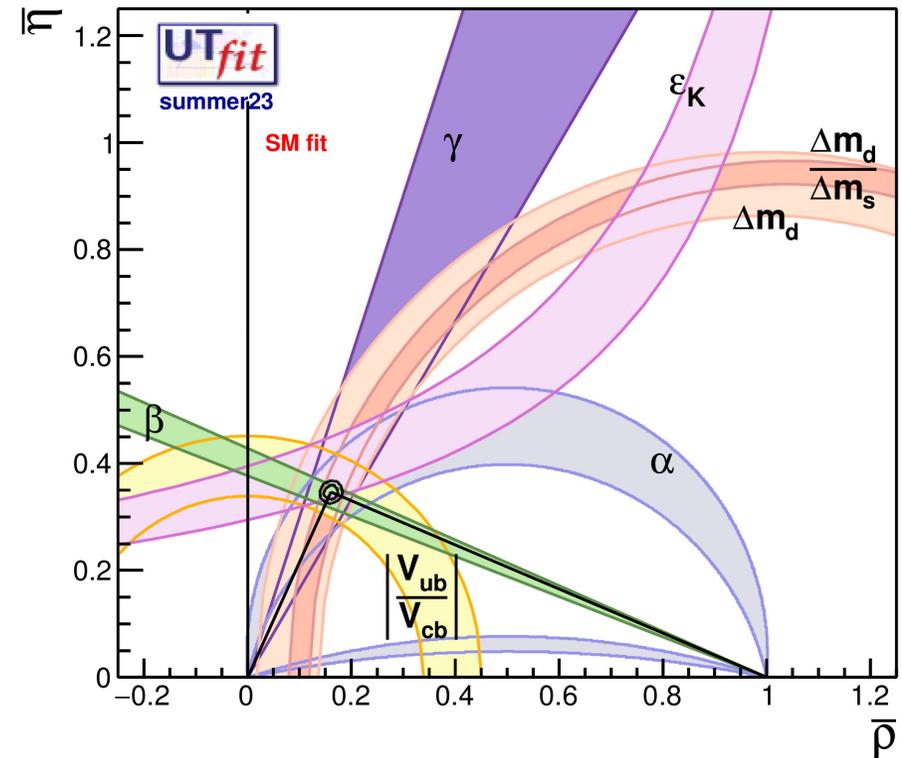
- Основана в 2014 году (совместно со Школой анализа данных Яндекса).
- Текущий состав: 4 старших исследователя, 6 аспирантов:
 - Д. А. Деркач, Ф. Д. Ратников, А. С. Болдырев, М. И. Гуцин
- НИУ ВШЭ – ассоциированный член LHCb с 2017 года.
- Также взаимодействие с экспериментами ЦЕРН:
 - SHiP
 - CMS
- В текущем составе группы в разное время:
 - LHCb CPV in Charm convener (Д. А. Деркач)
 - LHCb ML and Stat WG convener (Д. А. Деркач, М. И. Гуцин)
 - SHiP Muon Shield coordinator (Ф. Д. Ратников)



Интересы группы в ЦЕРН

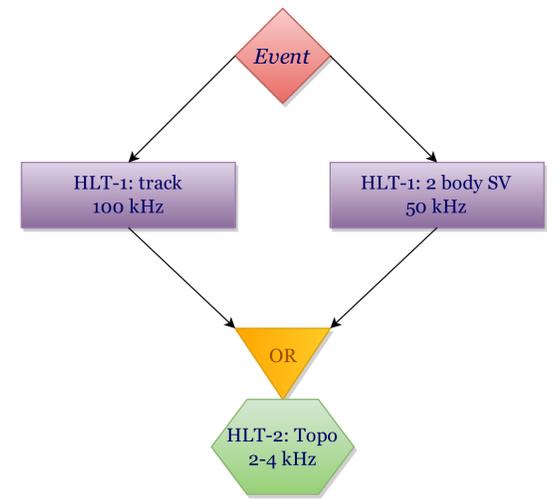
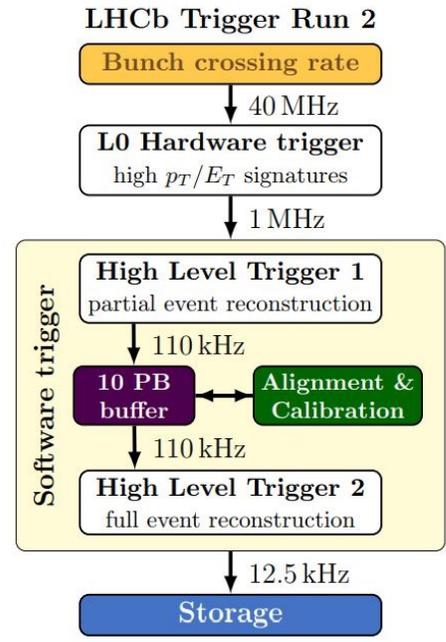
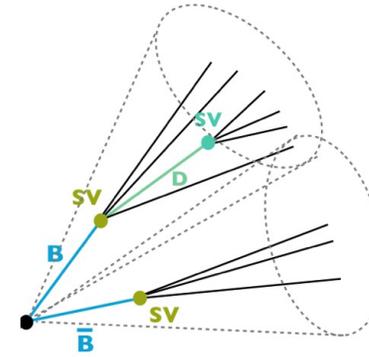
Разработка и применение методов машинного обучения

- Сбор экспериментальных данных в физике частиц:
 - Триггеры и онлайн обработка данных.
 - Идентификация заряженных и незаряженных частиц.
 - b-таггинг.
 - Сертификация данных.
 - Управление данными на суперкомпьютерах.
- Быстрая симуляция генеративными моделями.
- Оптимизация экспериментальных установок:
 - Калориметр LHCb
 - Магниты SHiP.
- Физический анализ данных: $B \rightarrow 4\mu, \Sigma \rightarrow p\mu\mu$.
- Феноменологические исследования в области слабых взаимодействий.



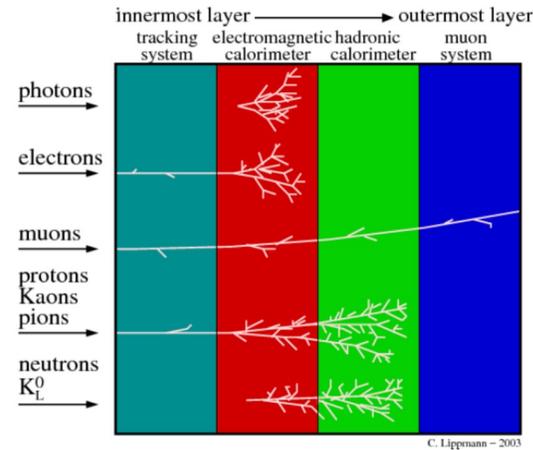
Топологический триггер LHCb

- Основан на частичной топологической реконструкции событий с возможностью утери некоторой информации.
- Работа в режиме онлайн в run 2 LHCb как один из основных триггеров.
- Основан на забывающих решающих деревьях.
- Ускорение за счёт пруннинга и компилируемого решения.
- Улучшение эффективности работы триггеров до 40% на разных распадах.

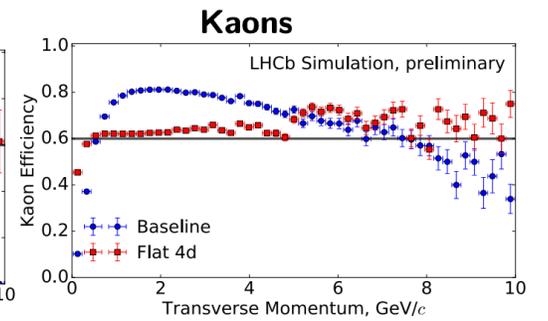
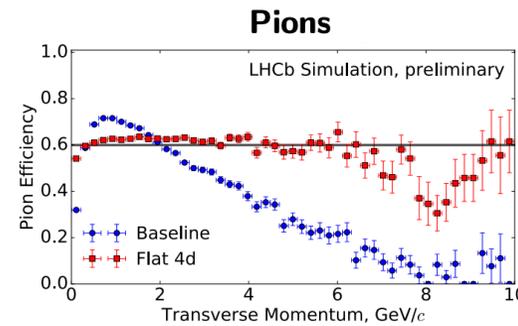
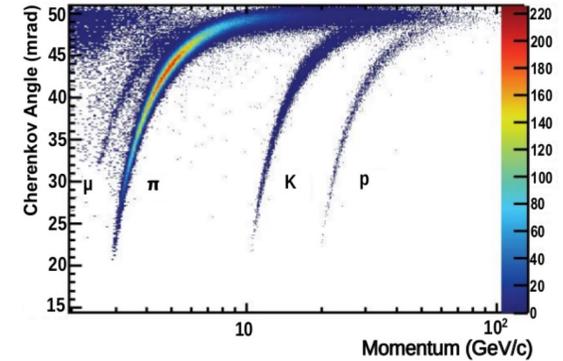


Идентификация частиц LHCb

- Комбинация информации со всех субдетекторов.
- Дополнительная очистка от шума и ghost частиц
- Внимание к характерным регионам и контроль профиля эффективности.
- Достигается за счёт изменения функции потерь при обучении нейросетей.



[Int J Mod Phys A30 (2015) 1530022]

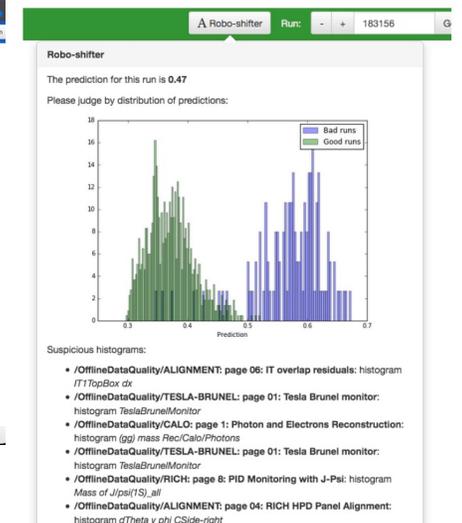
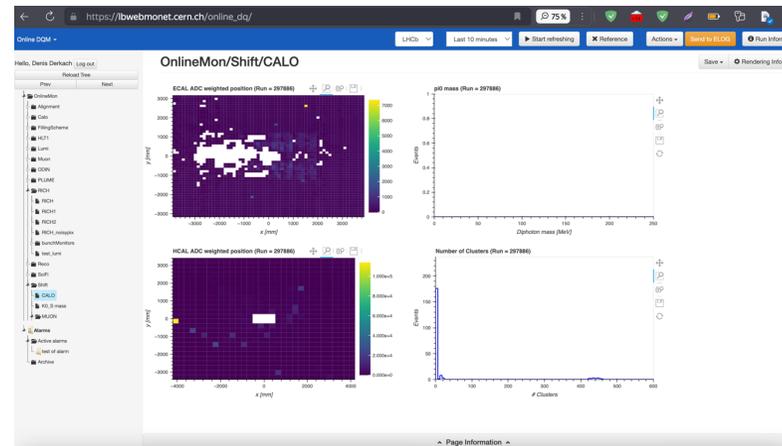
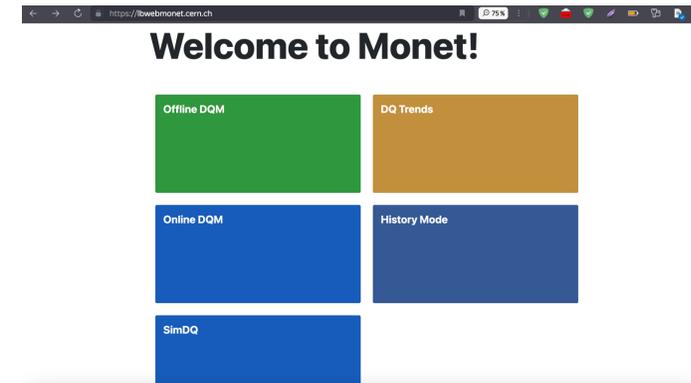


Flat4d, ProbNN



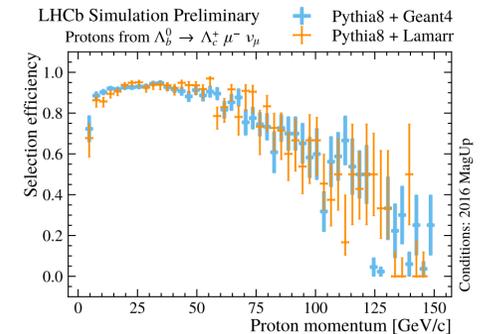
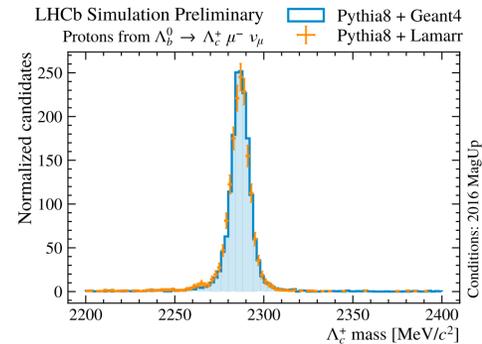
Система мониторинга Monet

- Базовый мониторинг LHCb с 2019 года и DUNE.
- Объединяет систему мониторинга данных LHCb:
 - Онлайн DQ мониторинг.
 - Оффлайн DQ мониторинг.
 - Экспертный мониторинг.
 - Мониторинг симуляции.
- Методы машинного обучения для быстрого анализа информации и привлечения внимания шифтера с быстрым анализом информации о причинах поломки.
- Веб-интерфейс для проведения мониторинга через Интернет.



Быстрая симуляция с помощью генеративных нейронных сетей

- Симуляция отклика субдетекторов, объединённая в единую библиотеку быстрой симуляции LHCb (Lamarr) или встроенную в Geant.
- Входные данные:
 - Детальная симуляция.
 - Реальные данные калибровочных каналов.
- Основной алгоритм: генеративно-состязательные сети.
 - Ускорения за счёт квантизации сетей.
 - Распределённая система тренировки.
- Текущий статус:
 - Тесты в реальных физических анализах.



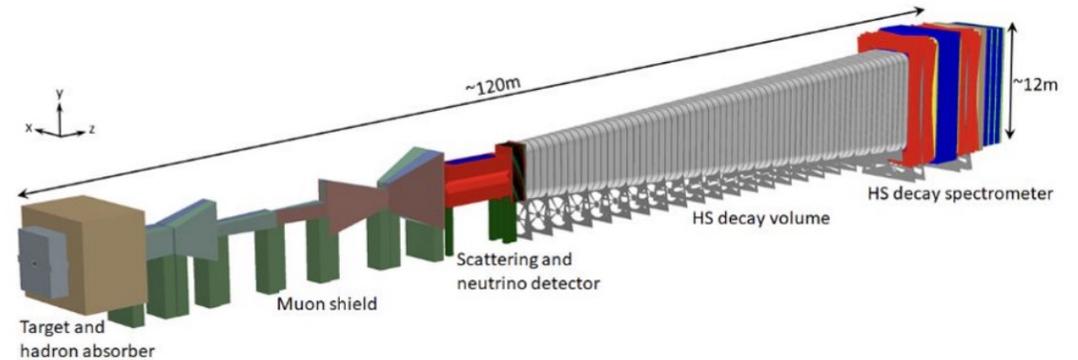
Detailed simulation: Pythia8 + Geant4
1M events @ 2.5 kHS06.s/event \approx 80 HS06.y

Ultra-fast simulation: Pythia8 + Lamarr
1M events @ 0.5 kHS06.s/event \approx 15 HS06.y

Ultra-fast simulation: Particle Gun + Lamarr
100M events @ 1 HS06.s/event \approx 4 HS06.y

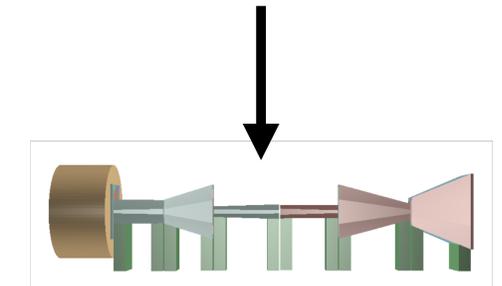
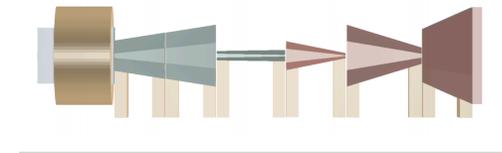
Оптимизация и дизайн экспериментальных установок

- Использование суррогатного моделирования методами (генеративного) машинного обучения для построения многопараметрической поверхности отклика.
- Переход к задаче оптимизации чёрного ящика для получения оптимальных параметров:
 - Глубокие гауссовские процессы.
- Гибкая целевая функция (цена эксперимента, эффективность каналов).
- Подход используется для оптимизации:
 - Мюонный щит SHiP.
 - Электромагнитный калориметр LHCb.



на **25%** снижена масса магнитов, на **~5 м** сокращена требуемая длина (экономия **~1 млн. CHF**)

Базовое решение



Оптимизированное с помощью цифрового двойника



Заключение

- Мы имеем обширный опыт в разработке и применении методов машинного обучения.
- Применяемые методы не «из коробки».
- Сотрудники лаборатории имеют смешанный опыт: прикладная математика и физика.
- Институт имеет экспертизу по самым современным направлениям методов искусственного интеллекта.

