

# Многомерная оптимизация детекторных систем на примере мюонной защиты в эксперименте SHiP

8-я Международная конференция по глубокому обучению в вычислительной физике  
DLCP 2024

Евгений Курбатов <sup>1,2</sup>  
Федр Ратников <sup>1,2</sup>  
Эдуард Урсов <sup>3,4</sup>

<sup>1</sup>ВШЭ  
<sup>2</sup>Яндекс ШАД  
<sup>3</sup>Humboldt-Universität zu Berlin  
<sup>4</sup>МИСИС



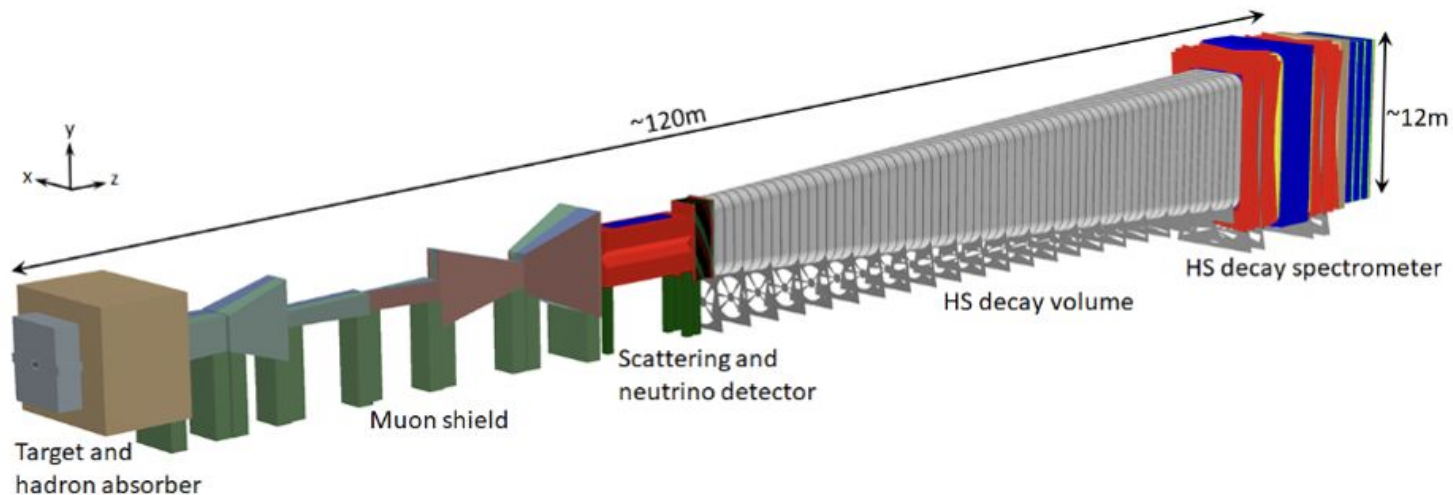
LAMBDA • HSE

Москва, 2024

# Введение

- В настоящее время одной из актуальных областей применения машинного обучения в физике высоких энергий является задача поиска оптимальных конфигураций детекторных систем. Целью такой оптимизации является нахождение баланса между способностью всех детекторов бесконфликтно выполнять свои задачи и стоимостью строительства или модернизации установки.
- Search for Hidden Particles (SHiP) является хорошим примером такого эксперимента.
- В докладе описываются подходы к комплексной оптимизации с использованием методов машинного обучения для детекторных систем в больших экспериментах на примере оптимизации мюонного щита в эксперименте SHiP.

# Эксперимент SHiP



- The Search for Hidden Particles (SHiP) это эксперимент нового поколения на SPS CERN
- Эксперимент с фиксированной мишенью
- $2 \times 10^{20}$  мишень-протонных взаимодействий за 5 лет,  $4 \times 10^{13}$  протонов в секунду

# Физические задачи эксперимента SHiP

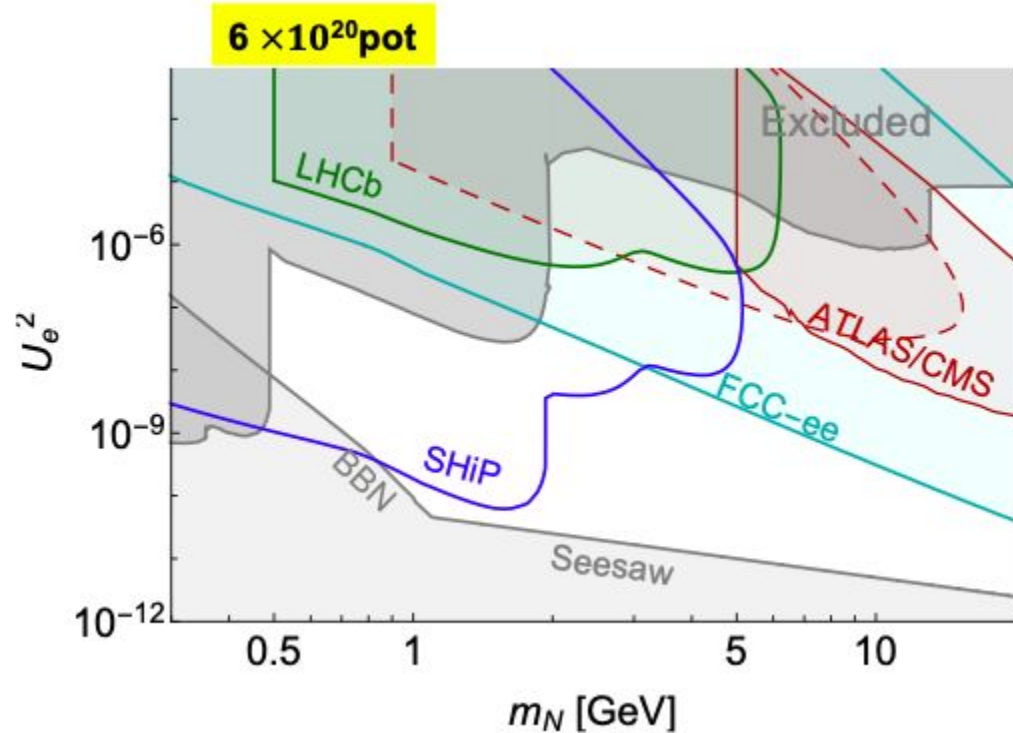
Основная задача - поиск heavy neutral lepton - HNL

Чувствительность SHiP к FIP на порядки выше, чем у конкурирующих проектов

Каналы распада:

$\text{HNL} \rightarrow \mu\pi$

$\text{HNL} \rightarrow \nu\mu$



# Физические задачи эксперимента SHiP

Огромная выборка тау-нейтрино может быть получена на SPS через распад

$$D_s \rightarrow \tau, \nu_\tau$$

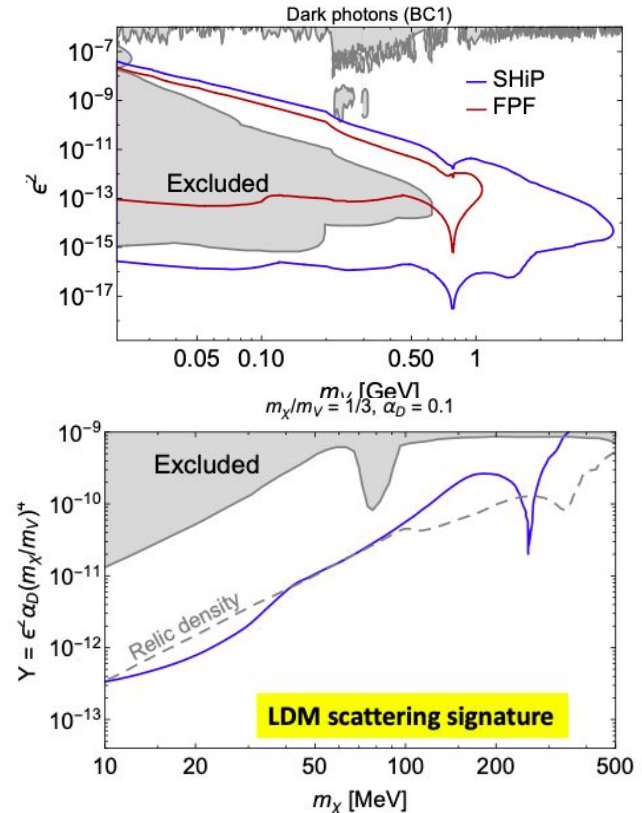
Важные тесты стандартной модели:

- Проверка Лептонной

Универсальности для нейтрино (точность 1-3% для отношений:  $n_e/n_\mu$ ,  $n_e/n_\tau$  и  $n_\mu/n_\tau$ )

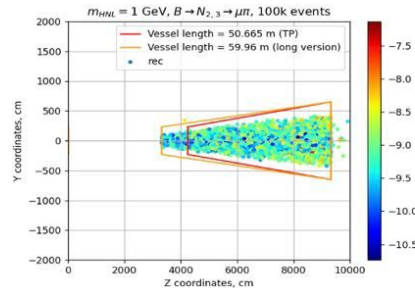
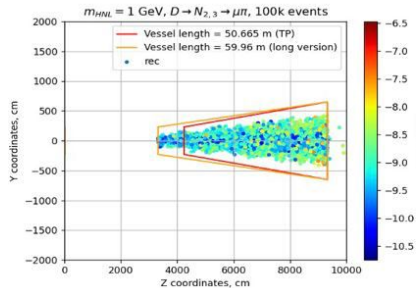
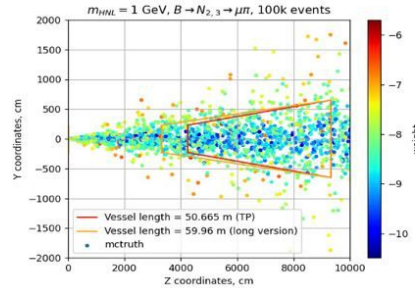
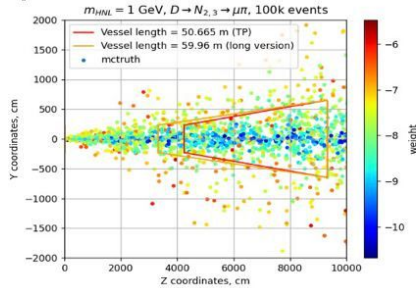
- Измерение сечений нейтрино до 100

ГэВ



# Влияние положения распадного объема на чувствительность к HNL

Spatial distribution of the HNL vertices. Color represents weight

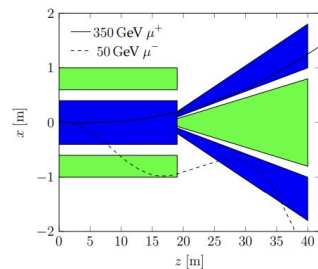
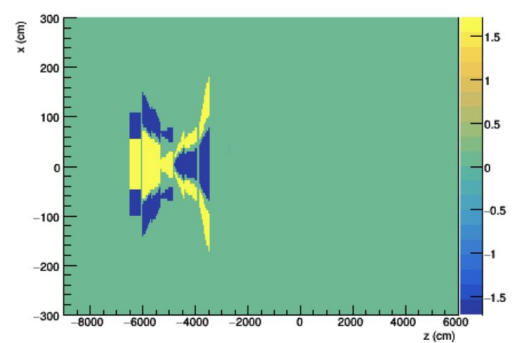


Ближе к мишени - больше чувствительность!

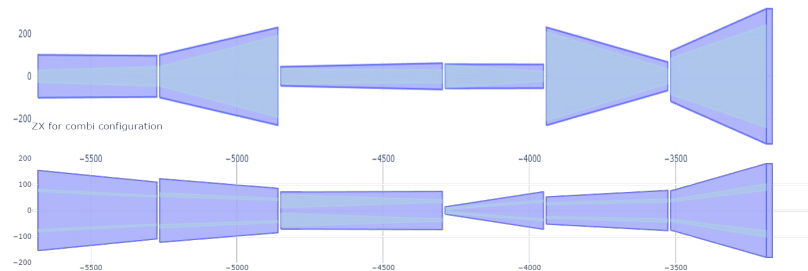
Color distribution strongly depends on gamma!

$$w = \frac{(L_2 - L_1)}{\beta \gamma c \tau} e^{-\frac{L_S}{\beta \gamma c \tau}}$$

# SHiP Muon Shield

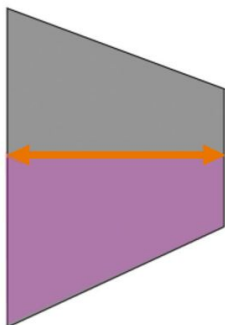
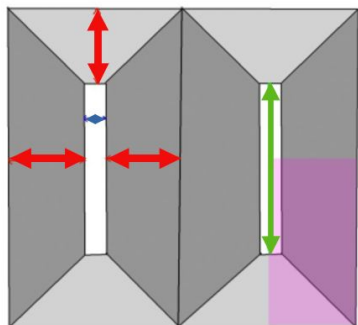


ZY for combi configuration



Front view

Side view

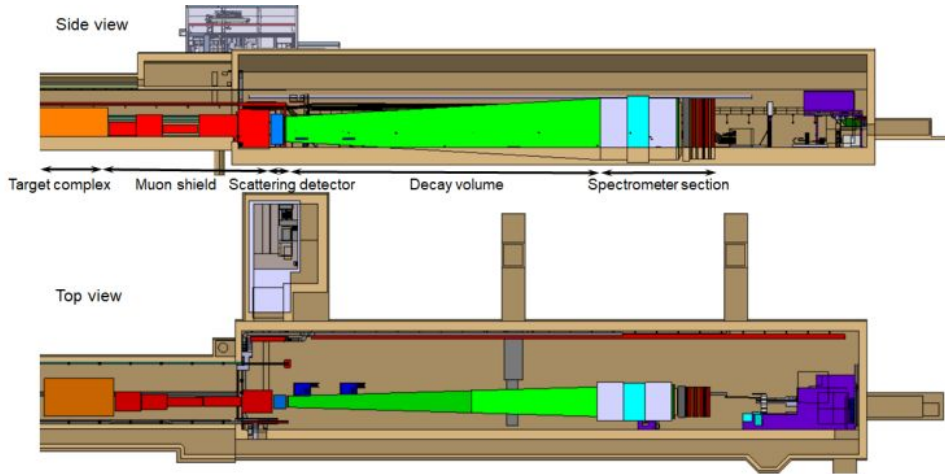


In

Out

- Один из ключевых элементов - мюонный щит, магнитная отклоняющая система длиной ~30 м, магнитное поле 1,7 Тл
- До 42 параметров, необходимых для определения конфигурации мюонного щита
- Качество экранирования является предметом споров и обсуждений нескольких научных групп
- Ожидаемое подавление потока мюонов ~6 порядков из  $10^{11}$  мюонов на спилл
- Процесс прохождения мюонов через защиту довольно стохастический

# Смена экспериментального зала, ECN3



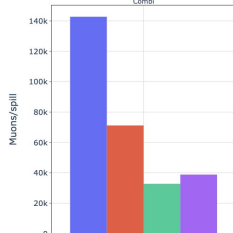
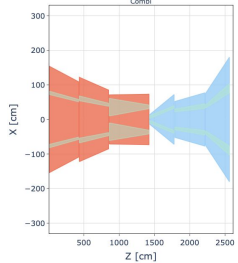
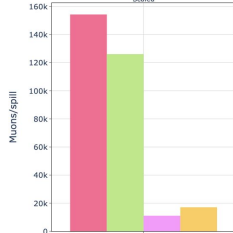
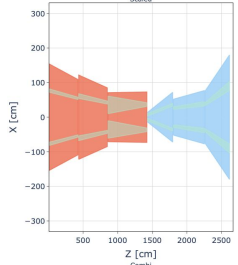
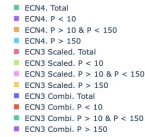
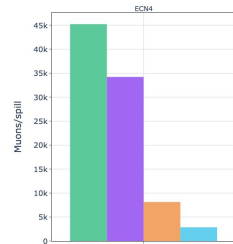
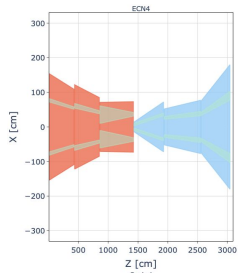
- ECN3 экспериментальный зал рассматривается как альтернативное место для проведения эксперимента
- Что накладывает новые ограничения на размеры мюонной защиты
- Старая конфигурация больше не является валидной
- Появилась необходимость укоротить защиту на 5 метров
- Изменились поперечные размеры трекингвых детекторов

Метрика качества мюонного щита - количество мюонов, прошедших щит и создавших не менее 3 попаданий из 4 станций слежения спектрометра для фиксированного сэмпла мюонов в объеме 1 спилла



# Попытка ручной оптимизации

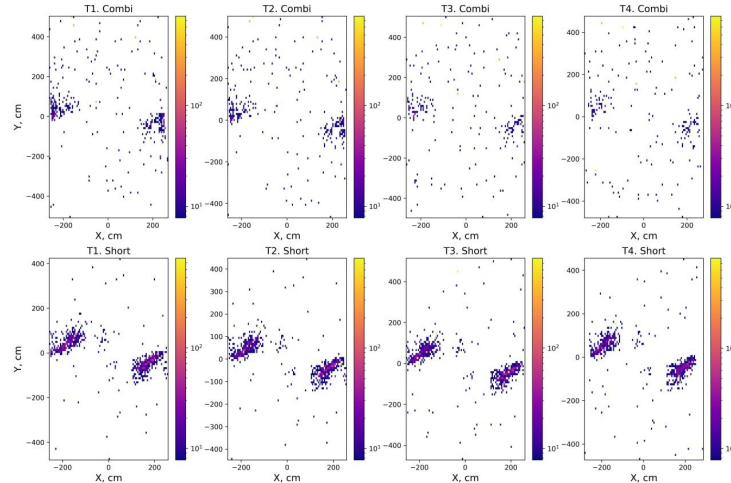
Muon track rates in Tracking Stations



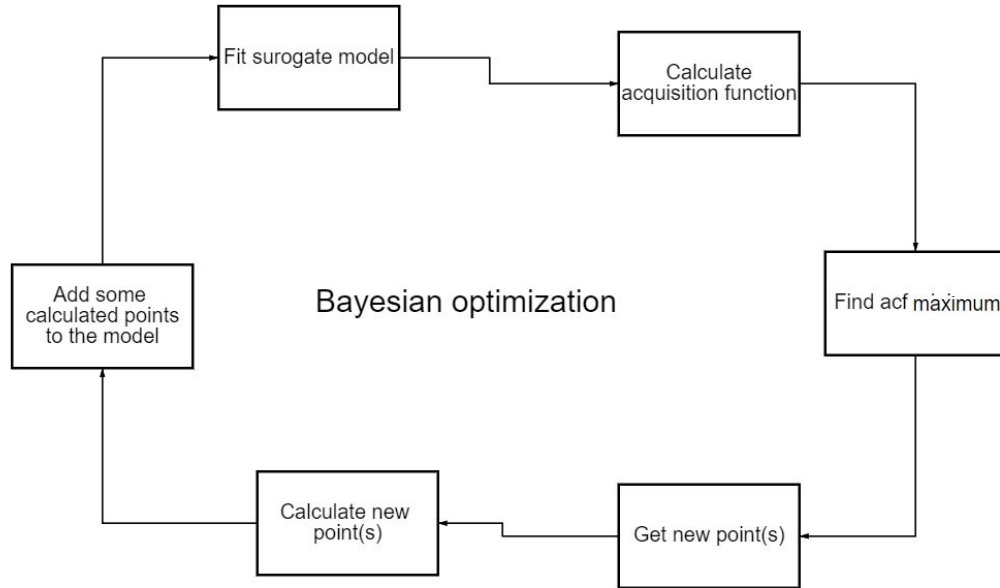
T1 & (T2 or T3) & T4

2 основные “ручные” версии Muon Shield:

- Равномерно сжатие по Z с 35 до 30м
- “Combi” - укорачивание только последних 3 магнитов



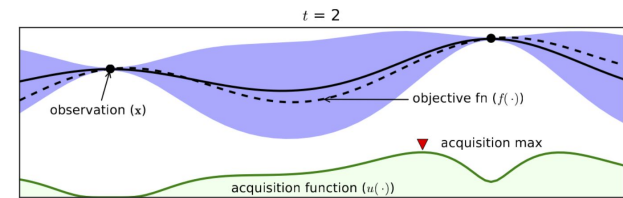
# Байесовская оптимизация



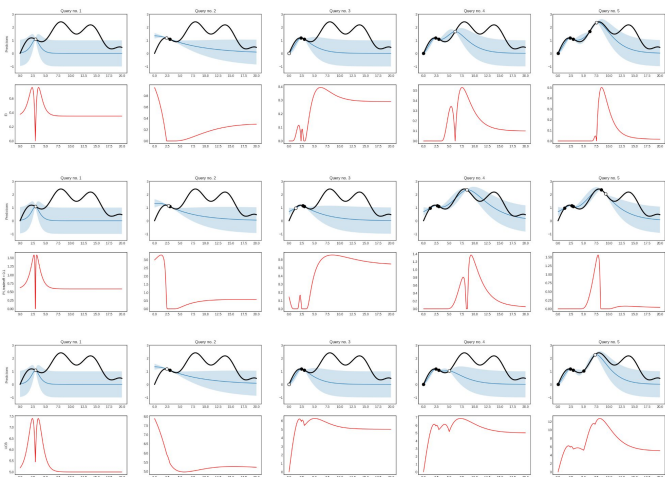
Глобальная оптимизация “черного ящика”

Подход дает прекрасные результаты, если:

- размерность  $< 10$
- непрерывная функция потерь
- форма пространства параметров не слишком сложна
- нет информации о градиенте для функции потерь



# Байесовская оптимизация, ACQF



Expected Improvement:  
 $EI(x) = E(\max(f(x) - \text{best}_f, 0))$

Probability of Improvement:  
 $qPI(X) = P(\max Y \geq \text{best}_f, Y \sim f(X), X = (x_1, \dots, x_q))$

Upper Confidence Bound  
 $UCB(x) = \mu(x) + \sqrt{\beta} * \sigma(x)$

# Байесовская оптимизация, суррогаты и ограничения

Вариант суррогата по умолчанию - гауссовский процесс (GP)

Оптимизация мюонного щита является ~40-мерной задачей -> необходимо проверить огромное количество конфигураций, чтобы найти глобальный минимум

- Вычислительная сложность  $O(n^3)$
- Потребление памяти  $O(n^2)$

Узкое горлышко ресурсов:

- Первые итерации: время на MC
- Время генерации новой точки
- Нехватка памяти для построения суррогата

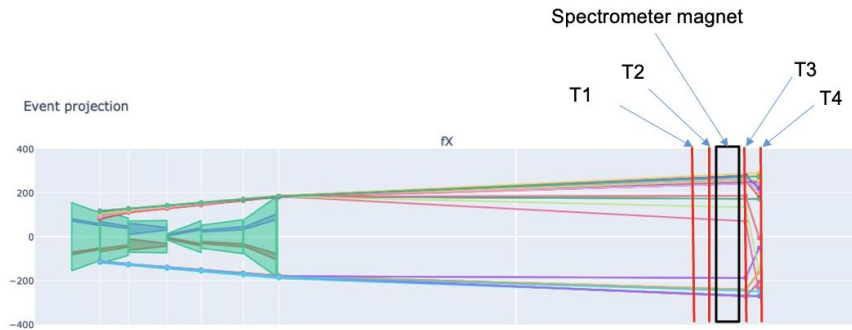
Базовый подход в оптимизации Muon Shield:

- Случайные точки
- High-dimensional Bayesian optimization with sparse axis-aligned subspaces (PMLR 161:493-503, 2021)
- Variational Nearest Neighbor Gaussian Process (<https://arxiv.org/abs/2202.01694>)

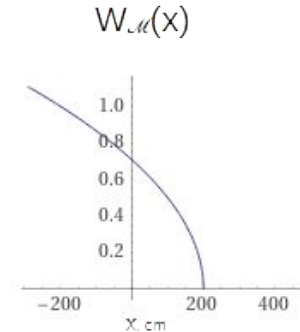


# Функция потерь

$$\left(1 + e^{\frac{10 \cdot (M - M^*)}{M^*}}\right) * \left(1 + \sum W_{\mu}\right)$$



$M$  - Вес мюонной защит  
 $M^*$  - константа  
 $W_{\mu}$  - Нелинейная функция  
координаты  $x$  для мюона в T1



Для МС используется специальный мюонный сэмпл:

- Убраны веса этапы генерации
- Размер уменьшен до минимально возможного, дающего стабильный результат

# Сэмпл. Форма. Скорость.

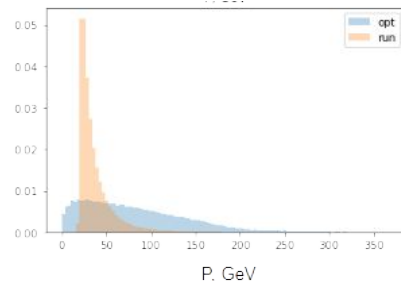
Большое число конфигураций, для которых необходимо провести симуляции накладывает ограничения на число событий, которые могут быть обработаны в рамках одной симуляции.

С одной стороны, слишком большое число событий существенно замедлит процесс оптимизации, т.к. больше времени будет уходить на каждую итерацию.

С другой стороны, при недостаточном числе событий, полученные оценки для целевой функции могут иметь слишком большую неопределенность.

Хорошо себя показал метод динамической подстройки: в случае нехватки данных для оценки целевой функции с заданной точностью запускалась повторная итерация симуляции с дальнейшим усреднением результатов, до достижения нужной точности.

Стоит отметить, что выбор состава сэмпла, а также формы импульсных распределений в значительной степени влияет как на результат оптимизации, так и на скорость симуляции.



# Сэмпл. Размер. Веса.

Еще одной проблемой работы с сэмплами с низким числом событий является проблема, возникающая при использовании в расчетах метрик весов этапа генерации.

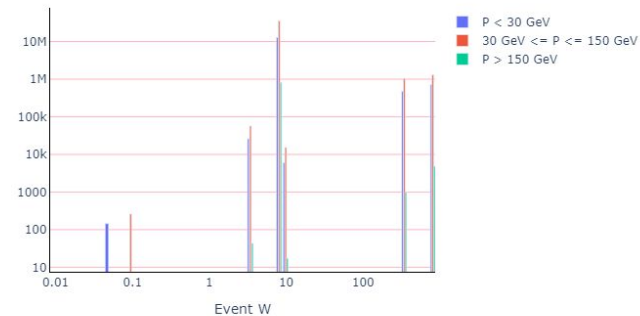
Если таковые используются, необходимо убедиться, что они не дестабилизируют результат.

Если такая проблема существует, то хорошим способом ее решить будет переход к невзвешенному сэмплу через процедуру “раскрутки” весов:

- из исходного сэмпла выбирается случайным образом событие (с вероятностью пропорциональной его весу) и добавляется в новый сэмпл
- в исходном сэмпле вес этого события уменьшается на минимальный из имеющихся в сэмпле весов
- процедура повторяется до достижения нужного числа событий в новом сэмпле

Данный подход позволил значительно снизить уровень неопределенности при работе с сэмплами ограниченных размеров.

Events Weights distribution

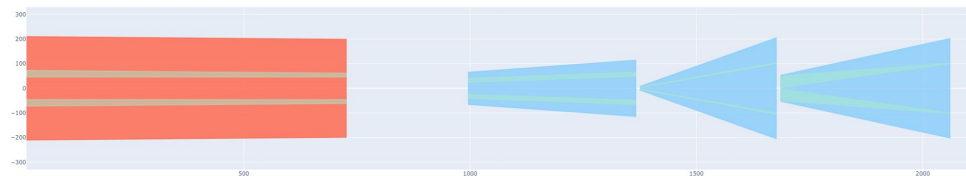
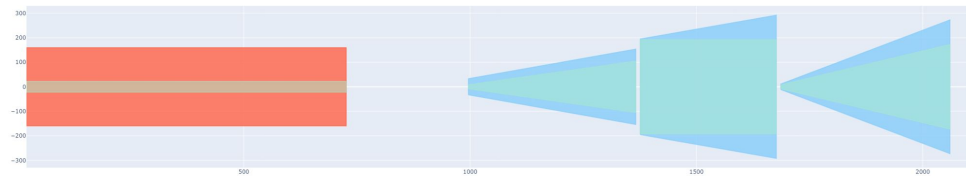


# Сверхпроводящая версия мюонной защиты

Идея: заменить первые 3 магнита щита одним коротким магнитом SC  $\sim 5T$ , чтобы сохранить интеграл магнитного поля прежним.

Оптимизационный цикл (за исключением FCN) оставался без изменений.

Количество параметров уменьшилось до 29 (параметром стал зазор между теплой и SC-частями)





# Новая функция потерь для SC защиты

Изначальная функция потерь была призвана оптимизировать поток мюонов через трекингтовую станцию. Однако, в дальнейшем появилась необходимость дополнить задачи:

- Минимизация потока через SND (сэндвич из металла и фотоэмульсии): в базовую функцию потерь был добавлен член регулирующий поток через SND
- Минимизация потока через veto-детектор, окружающий распадный объем: отдельная функция потерь, только контроль, оптимизация по функции из пункта выше

$$F = W * (T_{\text{flux}} + \text{SND}_{\text{flux}} + 1)$$

$$W = a_1 * W_{\text{sc}} + a_2 * W_{\text{warm}}$$

$$T_{\text{flux}} = a_3 * \max(0, (T - T_{\text{good}}))$$

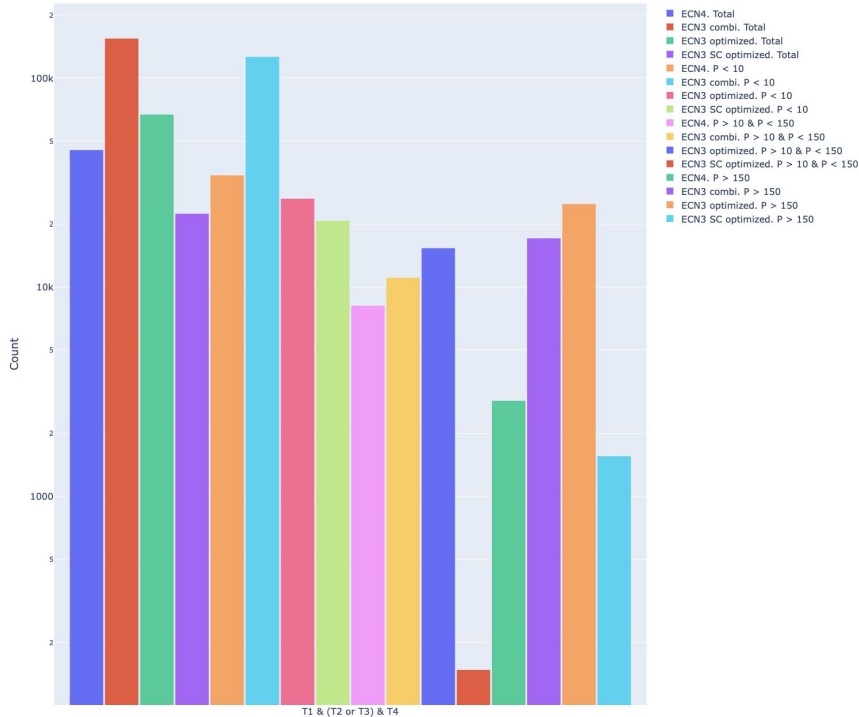
$$\text{SND}_{\text{flux}} = a_4 * \max(0, (S - S_{\text{good}}))$$

$a_i$  - параметры;  $T, S$  - функции потока мюонов через SND и T-станции, эквивалентны прошлым  $T_{\text{good}}, S_{\text{good}}$  - “достаточно хорошие” значения потоков, улучшать ниже нет смысла.



# Результаты оптимизации

Muon track rates in Tracking Stations



Shield	Rate [muons/spill]	Length [m]
ECN4	45k	35
ECN3 Combi	160k	30
ECN3 Optimized	67k	30
ECN3 SC Optimized	22k	20

# Выводы

- BDF/SHiP предоставляет четкую возможность обнаружить FIP в распадах тяжелых кварков (или закрыть эту "тему" экспериментально). Это дополнит поиски FIP на HL-LHC и будущем  $e^+e^-$  - коллайдере (где FIP можно искать в распадах бозонов)
- Muon Shield Bayesian Optimization превосходит ручную подстройку и обеспечивает хорошие результаты для мюонных защит SC/NC
- Полученное подавление потока мюонов удовлетворяют физическим требованиям
- Метод показал себя рабочим для решения нескольких независимых задач



# Backup slides

# Local Generative Surrogates

Проблема высокой размерности -> трудно найти точный глобальный минимум.

С другой стороны, у нас есть подход градиентного спуска для перехода к локальному минимуму.

Мы не можем использовать его из-за очень сложной формы целевой функции.

Не можем использовать его, потому что не можем передать градиент через MC-симуляцию.

Но если мы знаем хорошую начальную точку вблизи глобального минимума...

И если у нас есть хорошая генеративная суррогатная NN-модель для предсказания результатов моделирования для заданного набора параметров...

