

Генеративные модели для физики частиц

Ф. Ратников

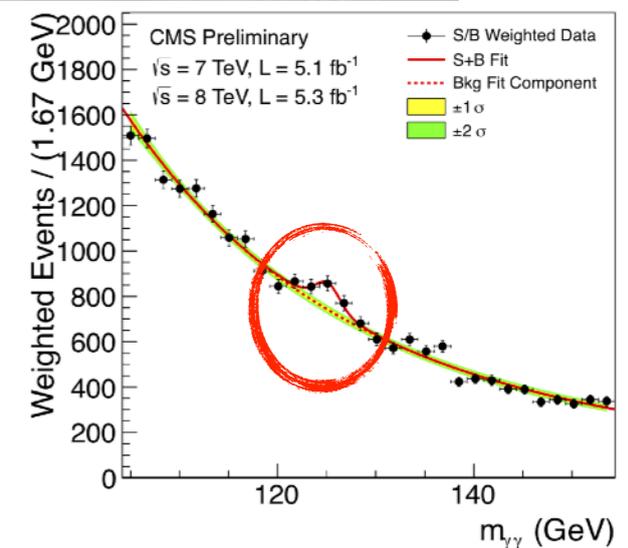
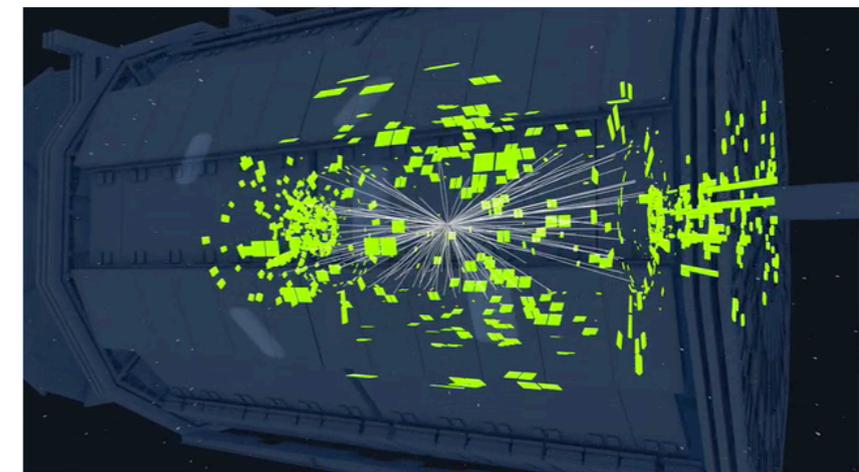
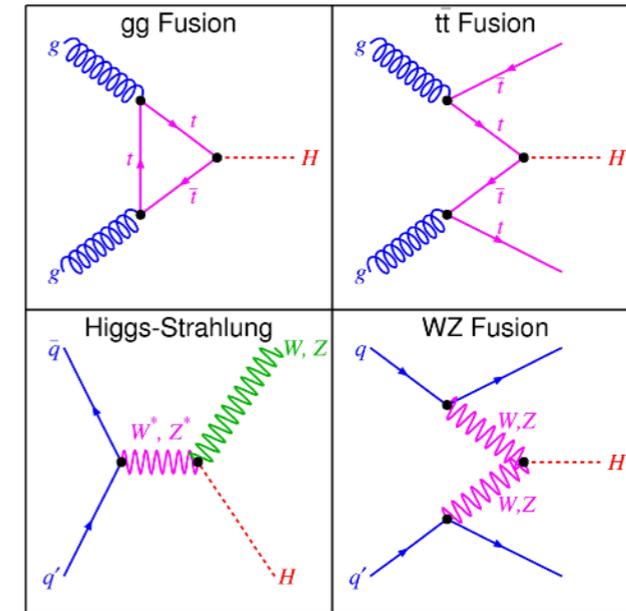
НИУ ВШЭ, лаборатория Лямбда

Санкт-Петербург, 28 августа 2023 г.



Длинная дорога данных ФВЭ

1. Взаимодействие сгустков частиц в ускорителе
2. Взаимодействие частиц с летящих навстречу или с частицами мишени
3. **Жесткое взаимодействие, описываемое моделью**
4. Образование партонов
5. Адронизация партонов в долгоживущие частицы
6. Взаимодействие частиц с веществом экспериментальной установки
7. Взаимодействие с детекторами частиц
8. Считывание данных с детекторов
9. Усиление и передача сигналов
10. Оцифровка сигналов
11. Объединение данных разных детекторов в единое событие
12. Триггерные отборы
13. Физический анализ
14. Измеренные наблюдаемые величины



Прямая и обратная задачи в ФВЭ

Прямая задача

- предполагаем некоторую, например Стандартную, модель элементарных частиц
- рассматриваем конкретный канал рождения частиц
- симулируем стохастические процессы и преобразования шагов 1..13
- оцениваем ожидаемый наблюдаемый эффект (14)
- сравниваем ожидаемый эффект с реально наблюдаемым в эксперименте

Обратная задача (Inference)

- рассматриваем конкретную параметрическую модель
 - › например, образование частиц темной материи некоторой массы

- делаем экспериментальное наблюдение
- проходим шаги 13..1
- оцениваем значения параметров модели

Физический анализ = решение обратной задачи на данных детектора

В условиях ФВЭ обратная задача слишком сложна, чтобы ее решать прямыми методами

- однако может быть решена итеративно через использование решения прямой задачи

Geant4

Основные компьютерные ресурсы ФВЭ тратятся на прецизионную симуляцию стохастических процессов на шаге (6)

Geant4 - универсальный пакет, реализующий шаги (6..8)

<https://geant4.web.cern.ch/>

несколько сотен человеко-лет разработки

более 1.5М линий кода

Большая часть из данных ФВЭ сгенерирована с помощью Geant4

(~ 5 Exabytes = 5M Tb)

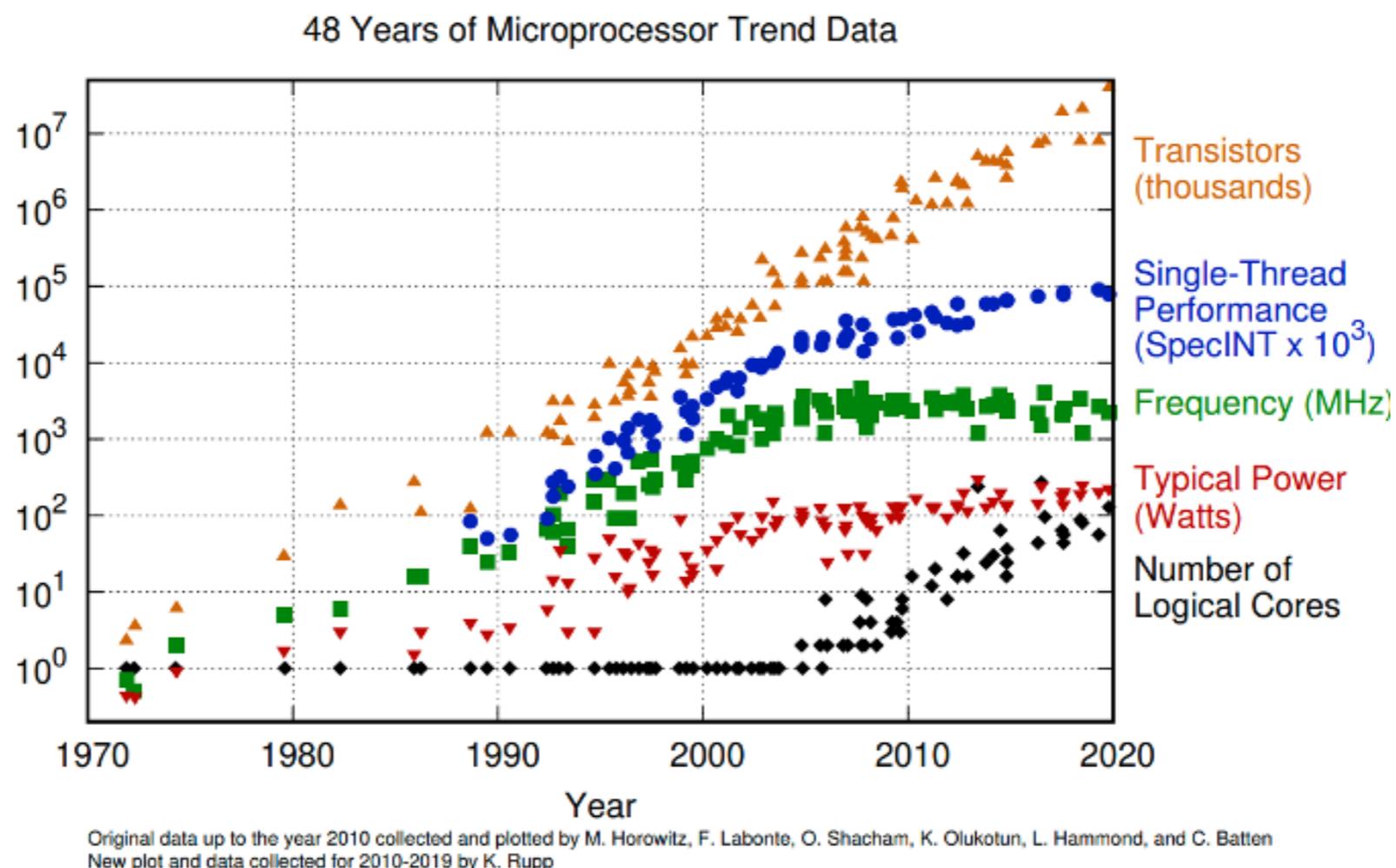
6. Взаимодействие частиц с веществом экспериментальной установки
7. Взаимодействие с детекторами частиц
8. Считывание данных с детекторов

The screenshot shows the Geant4 website homepage. At the top, there is a navigation menu with links for 'About', 'Download', 'Documentation', 'User Forum', 'Bug Reports', 'Events', and 'Contact Us'. The main header features the 'Geant4' logo and a descriptive text: 'Toolkit for the simulation of the passage of particles through matter. Its areas of application include high energy, nuclear and accelerator physics, as well as studies in medical and space science.' A 'Getting started' button is visible on the right. Below the header, the page is organized into several columns. The first column, 'Get started', includes a link 'I'm ready to start!'. The second column, 'Download', mentions 'Latest: 11.1.1'. The third column, 'Docs', has a link 'Read documentation'. The fourth column, 'News', lists several updates: '2023 Planned Features', '03 Mar 2023 Release 11.0.4', '10 Feb 2023 Release 11.1.1', '09 Dec 2022 Release 11.1', and '16 Sep 2022 Release 11.0.3'. Below the news are sections for 'About us' (with a link to 'What is Geant4, where it's used, details on Collaboration'), 'Collaboration' (with a link to 'Geant4 team and documents'), and 'Contribute' (with a link to 'How external users can contribute to Geant4'). At the bottom, there is an 'Events' section with a link to 'More' and a specific event: '9/26/2023 - 9/28/2023 Geant4 Training Course in Medicine 2023, Hokkaido University, Sapporo (Japan)'.

Генеративные модели в ФВЭ (FastSim)

Более 80% вычислительных ресурсов ФВЭ тратится на симуляцию данных экспериментов

Закон Мура вышел в насыщение



Требуются новые решения для симуляции данных высоких светимостей современных экспериментальных установок

Генеративные модели в ФВЭ (FastSim)

Более 80% вычислительных ресурсов ФВЭ тратится на симуляцию данных экспериментов

Закон Мура вышел в насыщение

требуются новые решения для симуляции данных высоких светимостей

Симуляция GEANT - переходная функция между (немногими) параметрами частицы и (макроскопическим) откликом детектора

вычисляемая на микроскопическом уровне детализации

Идея: натренировать (более простую и быструю) суррогатную модель для воспроизведения этой не многомерной переходной функции

Генеративные модели для физических симуляций

Самая эффективная генеративная модель, обученная на имеющемся семпле - “ближайший сосед”

выбираем событие обучающей выборки, ближайшее к требуемому в пространстве условий

Проблемы:

- › дискретные генерируемые объекты
- › нарушение инвариантов между условиями и свойствами объектами

хотелось бы интерполировать между несколькими тренировочными объектами

- › сделать распределения непрерывными
- › приблизить инварианты

Прямая интерполяция объектов



Результатом является арифметическое смешивание

Параметрическая интерполяция

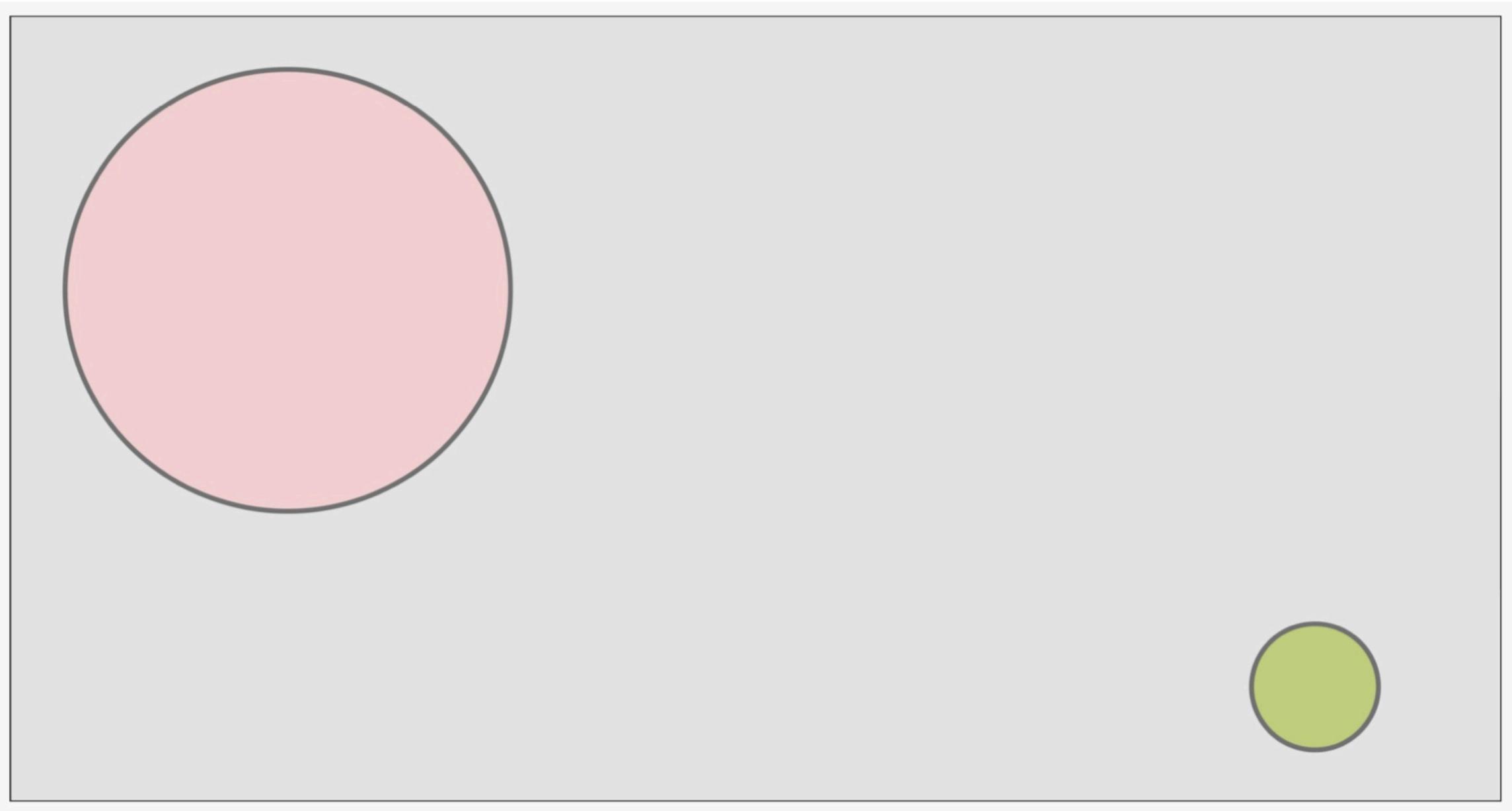
Смешивание представлений

Пара объектов представляются двумя точками в латентном пространстве

- ▶ можем соединить точки линией, проходящей через заселённые области фазового объёма латентных переменных

каждой точке на линии соответствует некоторый объект
объекты переходят один в другой непрерывно

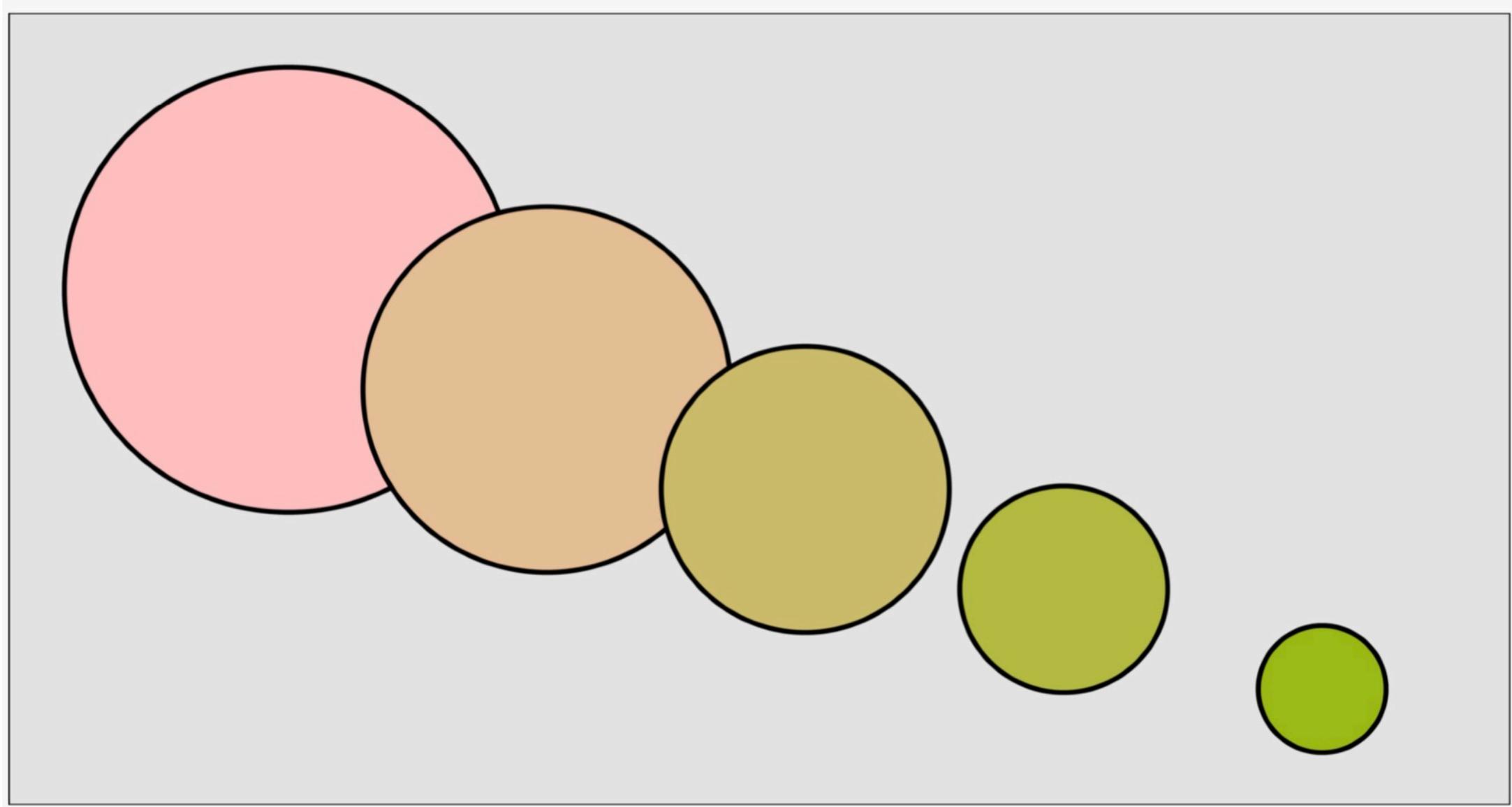
Параметрическая интерполяция



Латентное представление:

- ▶ положение центра
- ▶ радиус
- ▶ цвет

Параметрическая интерполяция

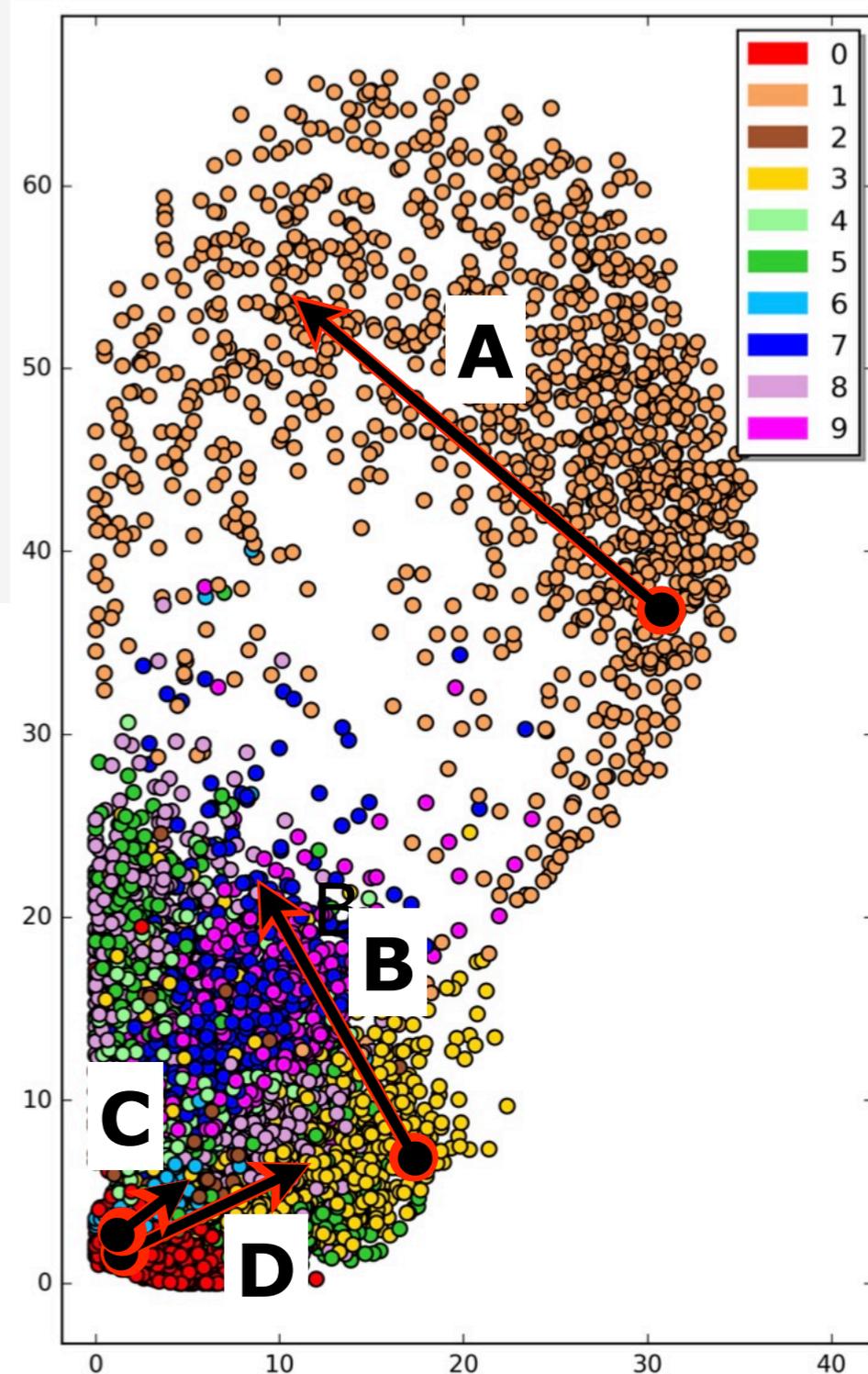


Промежуточные состояния являются также валидными объектами

Параметрическое смешивание



Промежуточные состояния являются также валидными объектами



Интерполяция объектов

Близкая арифметическая интерполяция работает для физических объектов:

если $\mathbf{c}_1 \rightarrow \mathbf{f}_1$, $\mathbf{c}_2 \rightarrow \mathbf{f}_2$ и $\mathbf{c} = \alpha\mathbf{c}_1 + (1 - \alpha)\mathbf{c}_2$, то
 $\mathbf{c} \rightarrow \mathbf{f} = \alpha\mathbf{f}_1 + (1 - \alpha)\mathbf{f}_2$

... но не работает для вариативностей:

$\sigma(\mathbf{f}) = \alpha^2\sigma(\mathbf{f}_1) + (1 - \alpha)^2\sigma(\mathbf{f}_2) \neq \alpha\sigma(\mathbf{f}_1) + (1 - \alpha)\sigma(\mathbf{f}_2)$

Требуется параметрическая интерполяция

требуется адекватная генеративная модель

Суррогатные модели для физических симуляций

Зачем нужна быстрая симуляция?

- нагенерить много данных за разумное время

Зачем нужно много данных?

- отследить тонкие или редкие эффекты в данных

- › эти тонкие и редкие эффекты требуют большого количества информации

- суррогатная модель *ab initio* содержит информацию из тренировочного сэмпла

- › тренировочный сэмпл должен быть достаточно большим

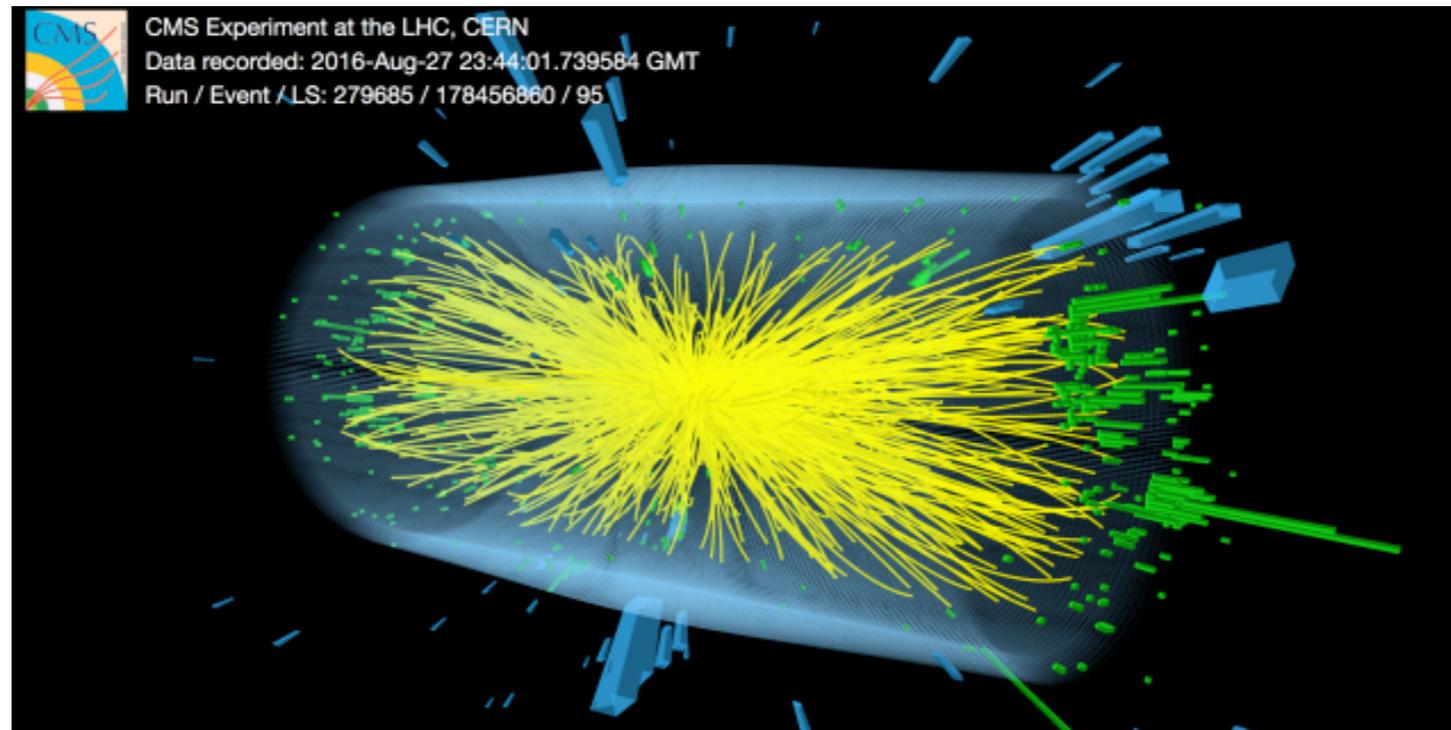
Для обучения суррогатной модели потребуется **больше данных**, чем количество данных, которые нам **требуется нагенерить для достижения необходимой точности**

Суррогатные модели для физических симуляций

Глобальная суррогатная модель непрактична

Суррогатная модель осмысленна, если:

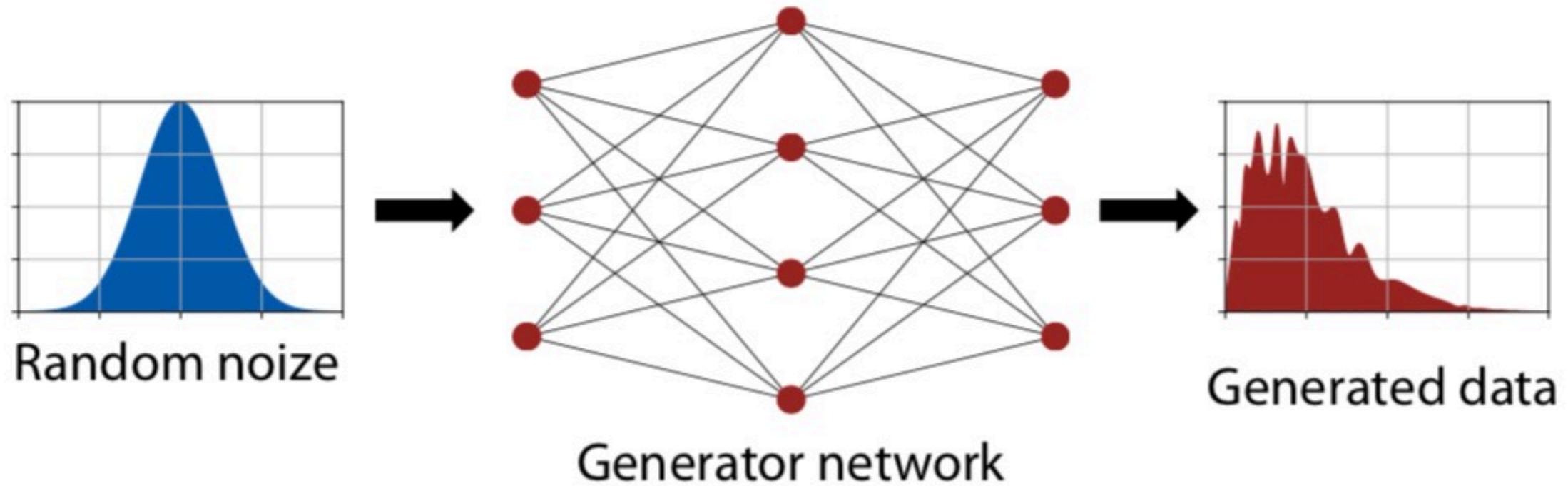
- генерация объектов может быть факторизована
- точность генерируемой единицы ограничена небольшим количеством информации
 - несложная модель натренированная на небольшом датасете



несколько сотен частиц в одном столкновении на БАК

- каждая частица может симулироваться независимо

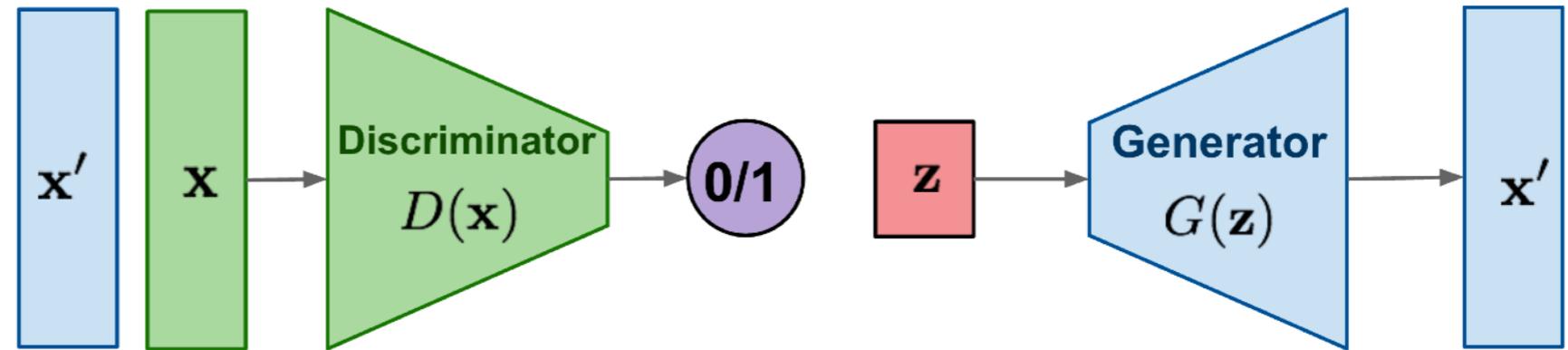
Генеративная модель ML



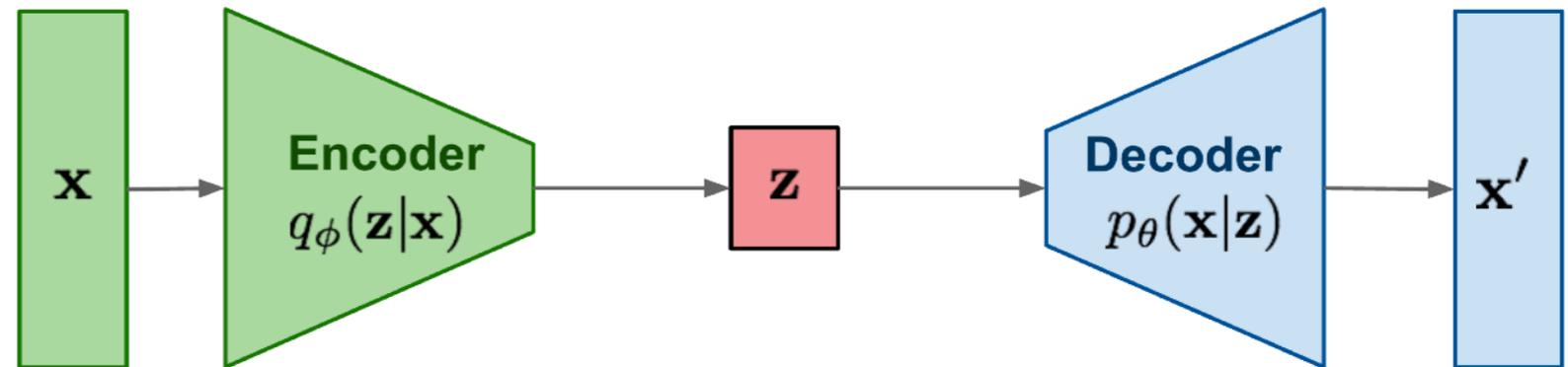
Обучаем переходную функцию из стандартного распределения в требуемое распределение

Виды (обучения) генеративных моделей

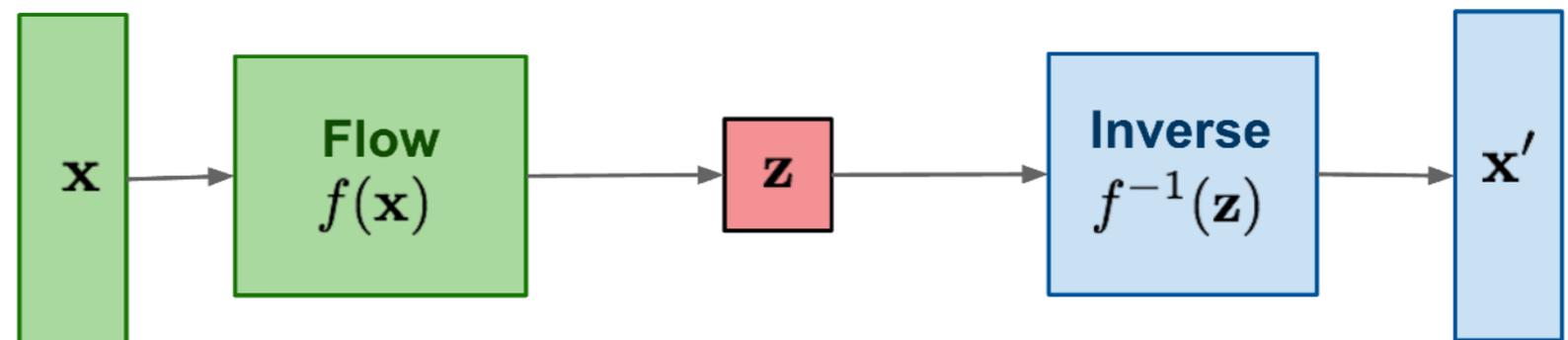
GAN: Adversarial training



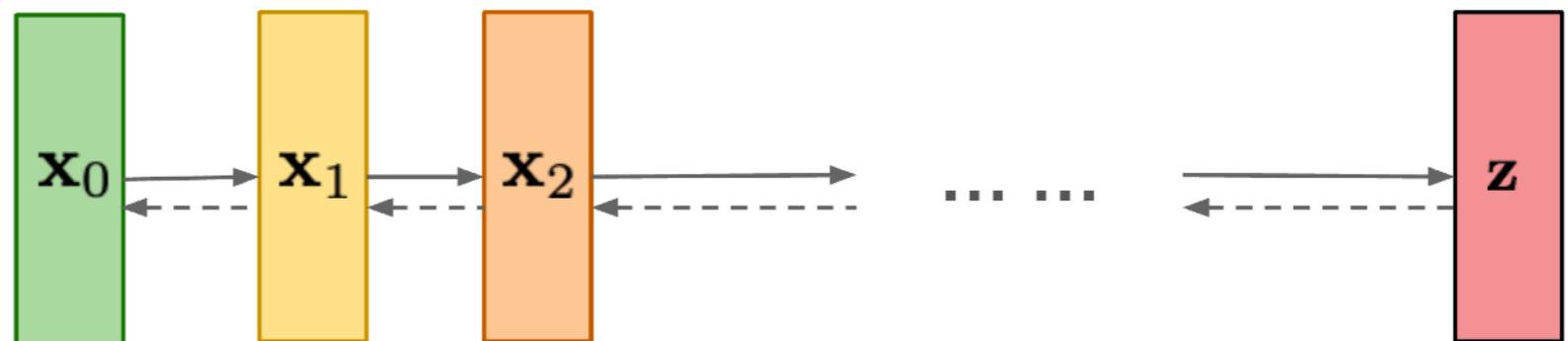
VAE: maximize variational lower bound



Flow-based models: Invertible transform of distributions

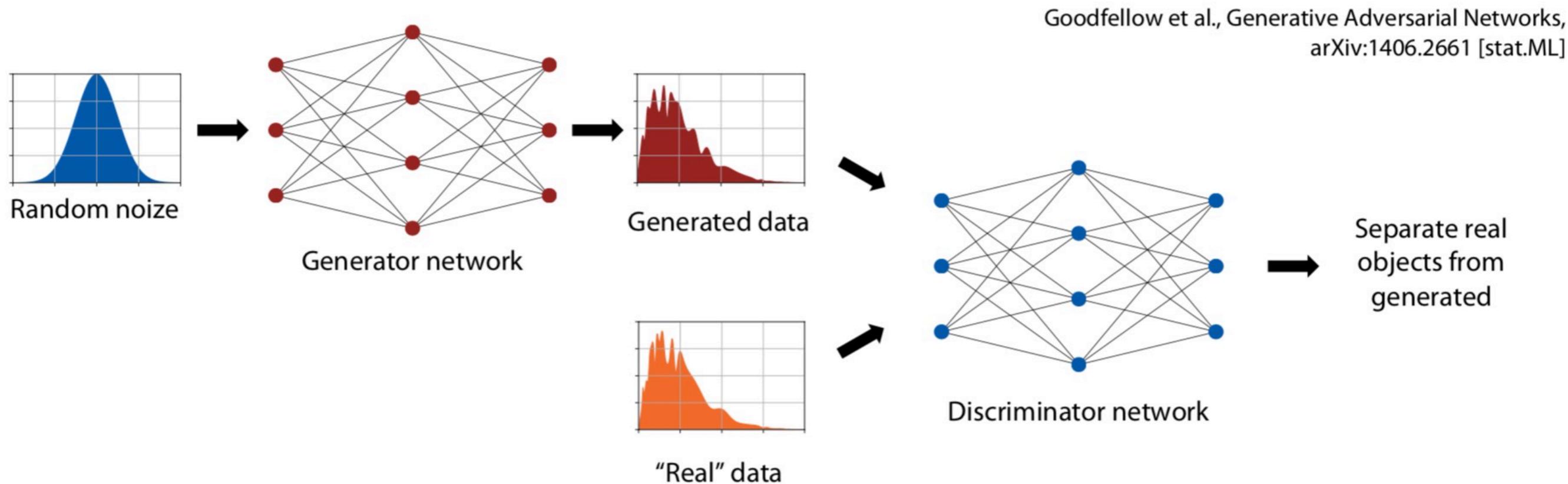


Diffusion models: Gradually add Gaussian noise and then reverse



Практика

Генеративно-сопоставительные сети



Метрика соответствия: насколько хорошо другая нейронная сеть может различить генерированные и реальные объекты

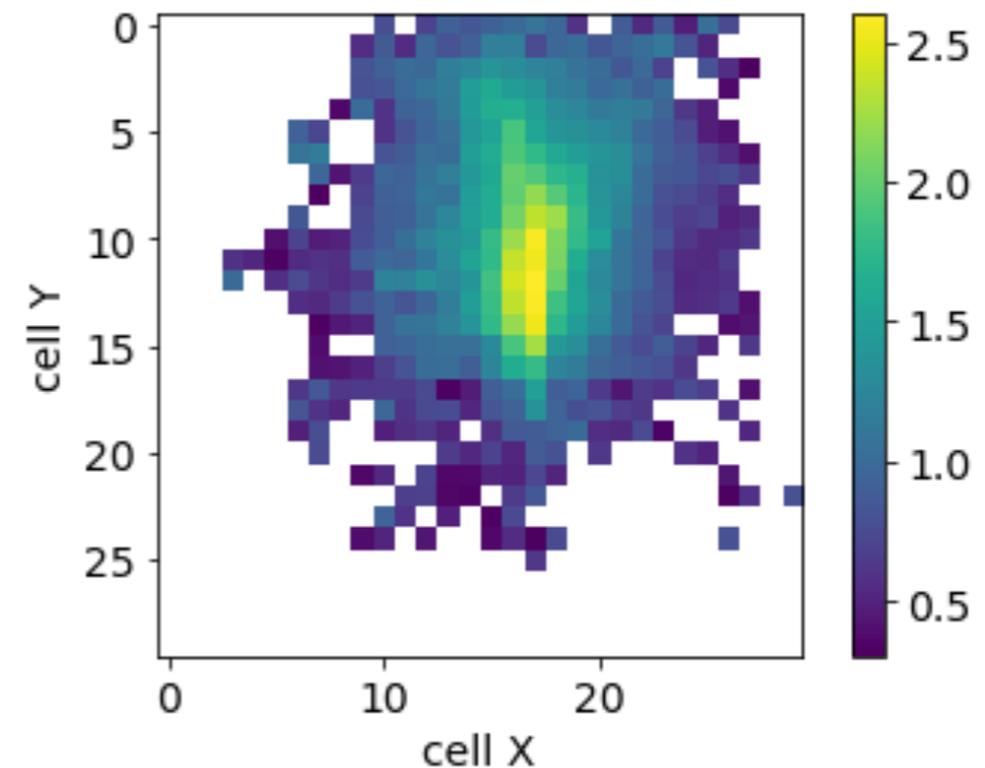
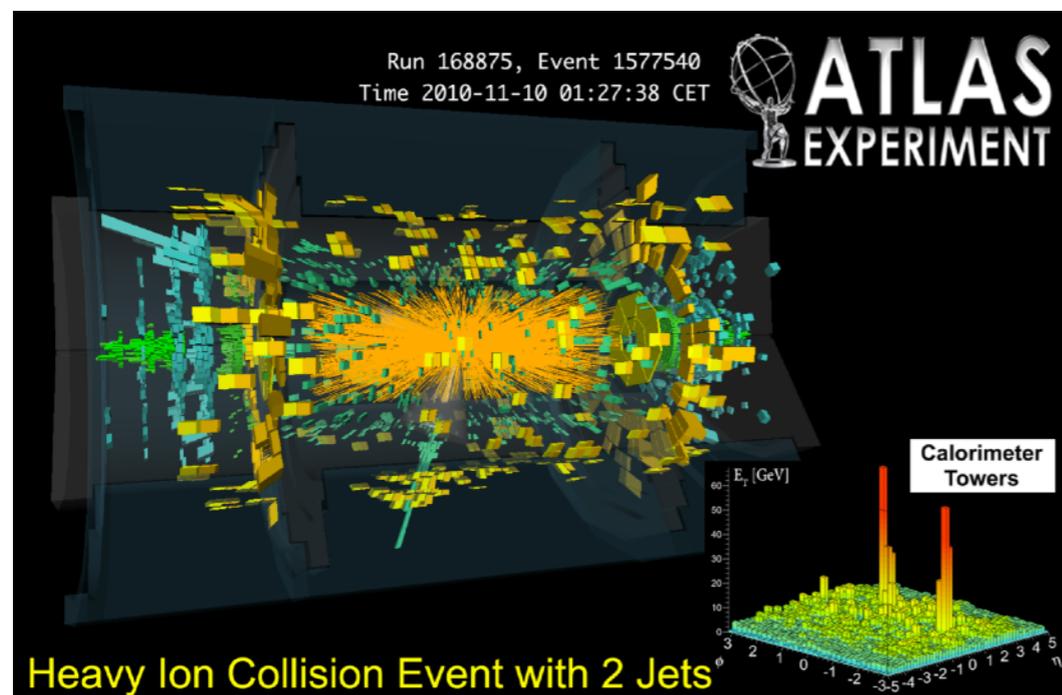
Генеративные модели для калориметров

Калориметры ФВЭ обычно имеют матричную структуру

- 2D (X,Y), 3D (X,Y,Z), 4D (X,Y,Z,время), ...

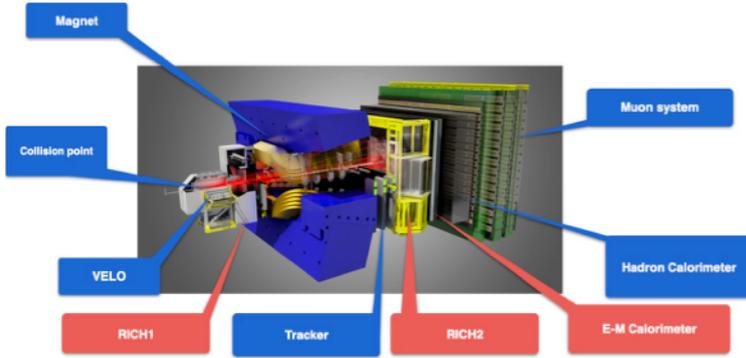
В каждой ячейке выделяется некоторая энергия

- отклик калориметра аналогичен картинке



Генеративные модели для калориметров

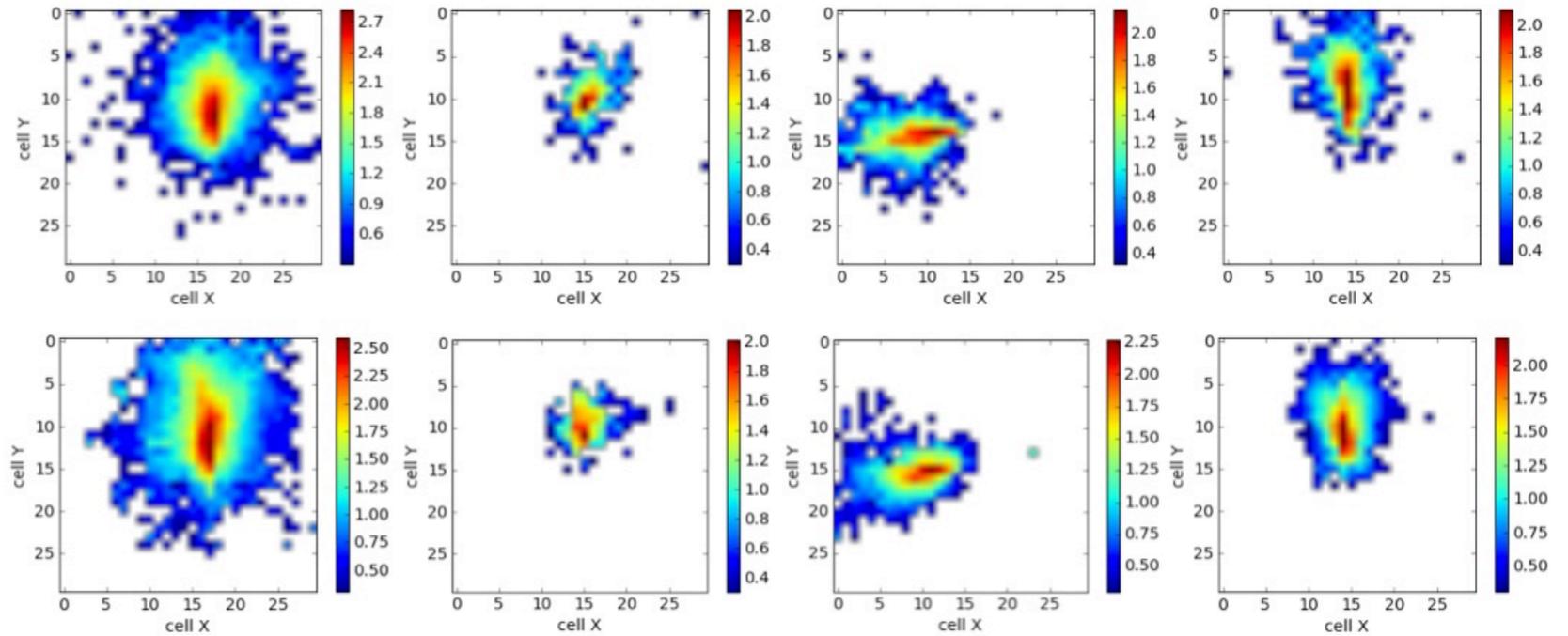
arXiv:1812.01319



GEANT Simulated

$\log_{10}(\text{cell energy})$

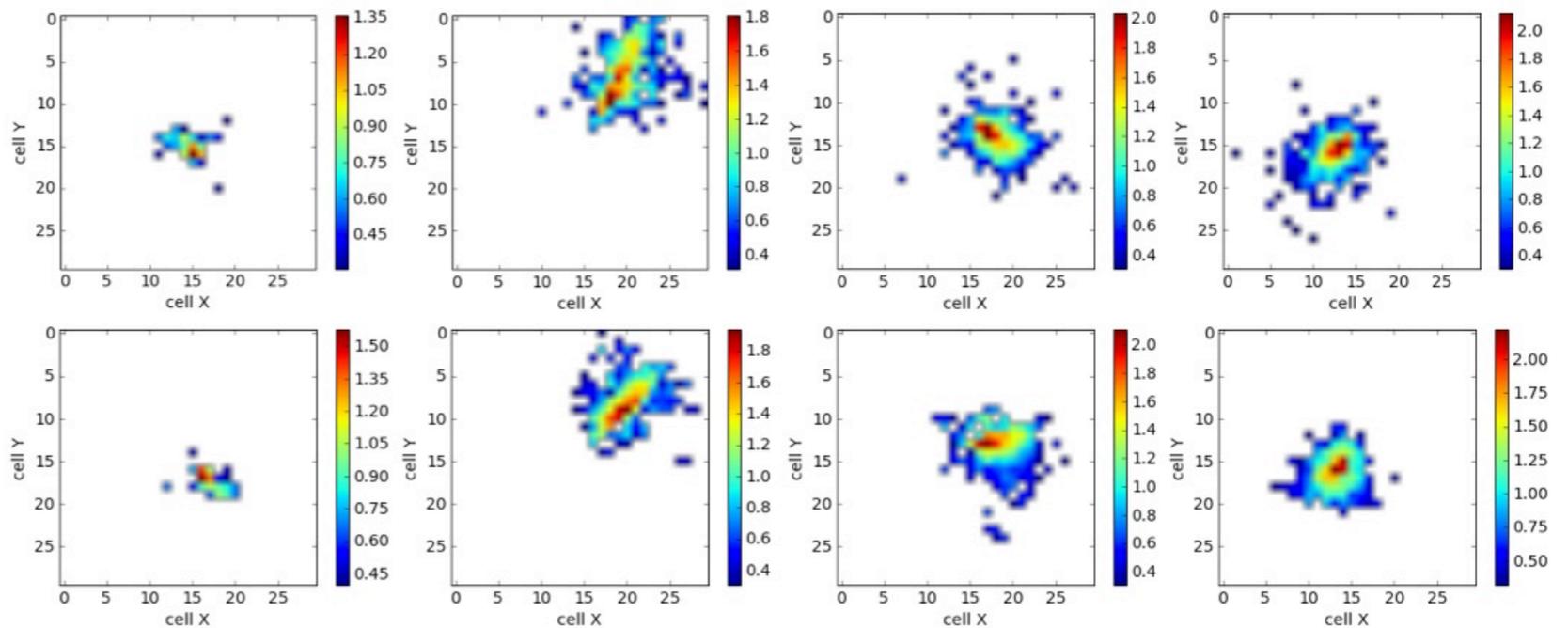
GAN Generated



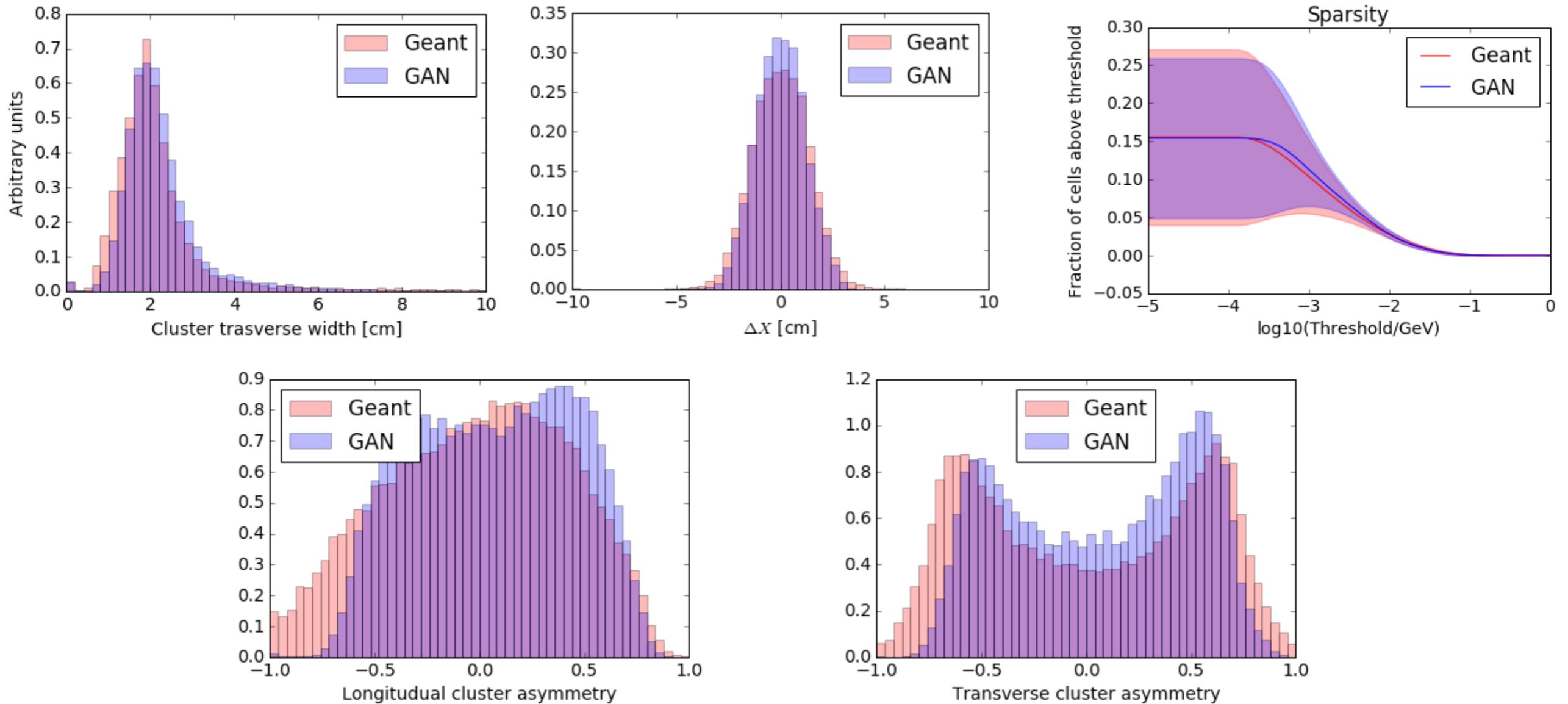
GEANT Simulated

$\log_{10}(\text{cell energy})$

GAN Generated



Генеративные модели



Сгенерировать выглядящие похоже отклики - несложно

Сложно воспроизвести маргинальные распределения

особенно, если их список не известен a priori

Приоритизация обучения

Генеративные модели неидеальны

- процесс обучения в целом агностичен по отношению к физическим метрикам

 - › некоторые воспроизводятся хорошо, некоторые совсем нет

- для научного использования качество по некоторым конкретным физическим метрикам является приоритетным

Как добиться приоритетного обучения распределениям приоритетных метрик?

- если метрика выражается через вычислительный граф - вставить в дискриминатор

 - › если метрика сложна?

Суррогатный регрессор

Если требуемая метрика сложна, заменим ее на суррогатный регрессор, обученный на этой метрике

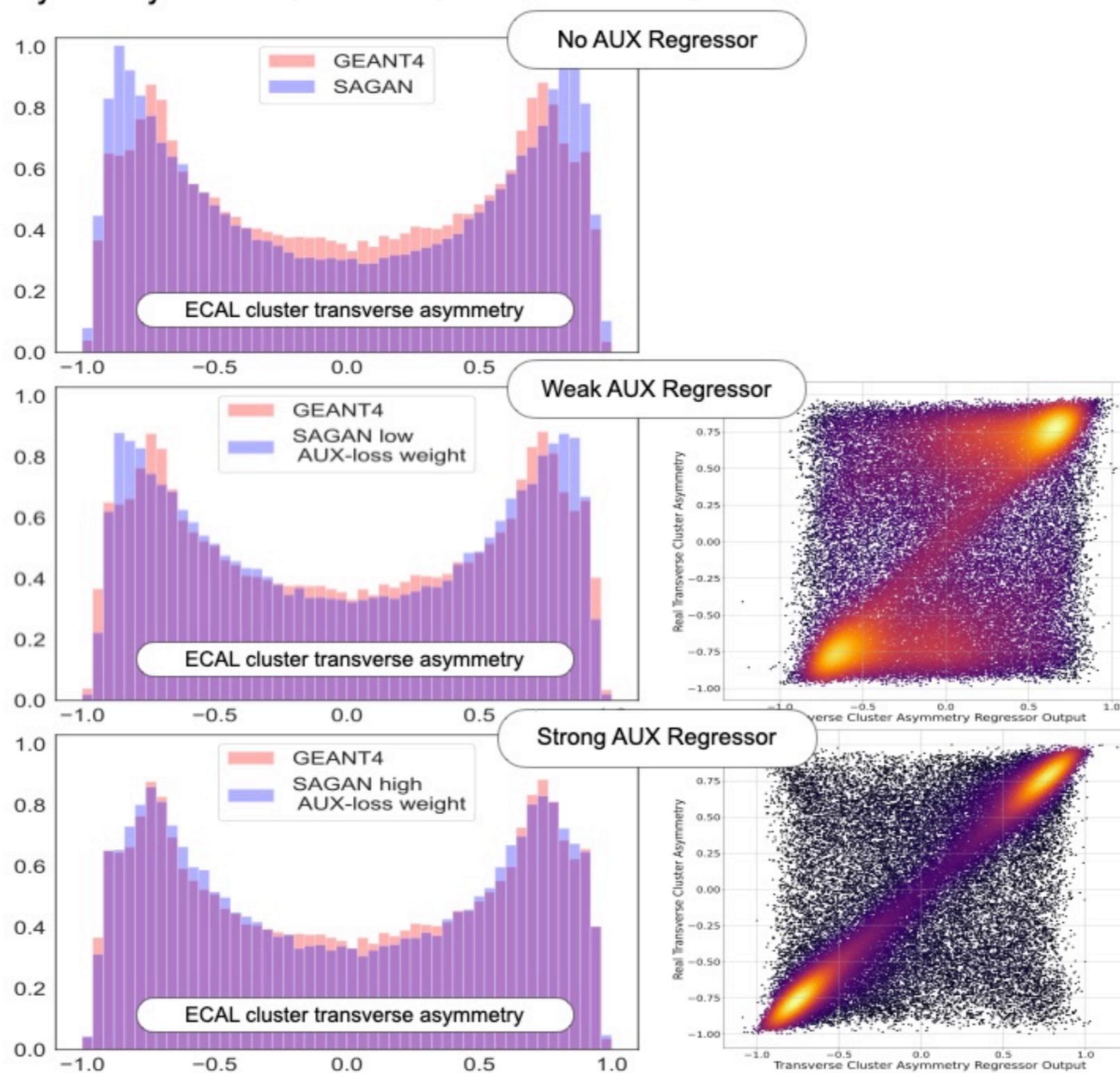
- обучим генеративную модель с использованием суррогатного регрессора

- для обеспечения хорошего качества воспроизведения требуемой метрики, воспроизведение метрики суррогатным регрессором не обязательно должно быть хорошим

- › регрессор применяется как к генерируемым, так и к истинным объектам

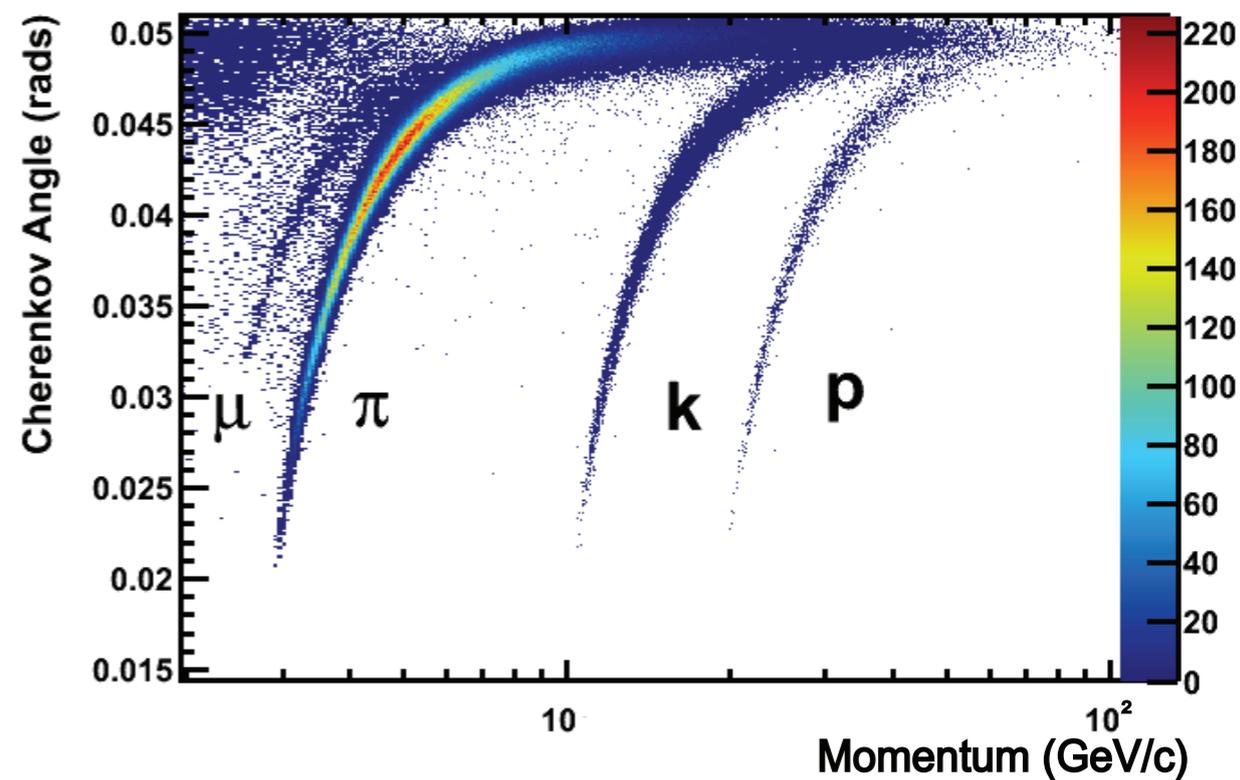
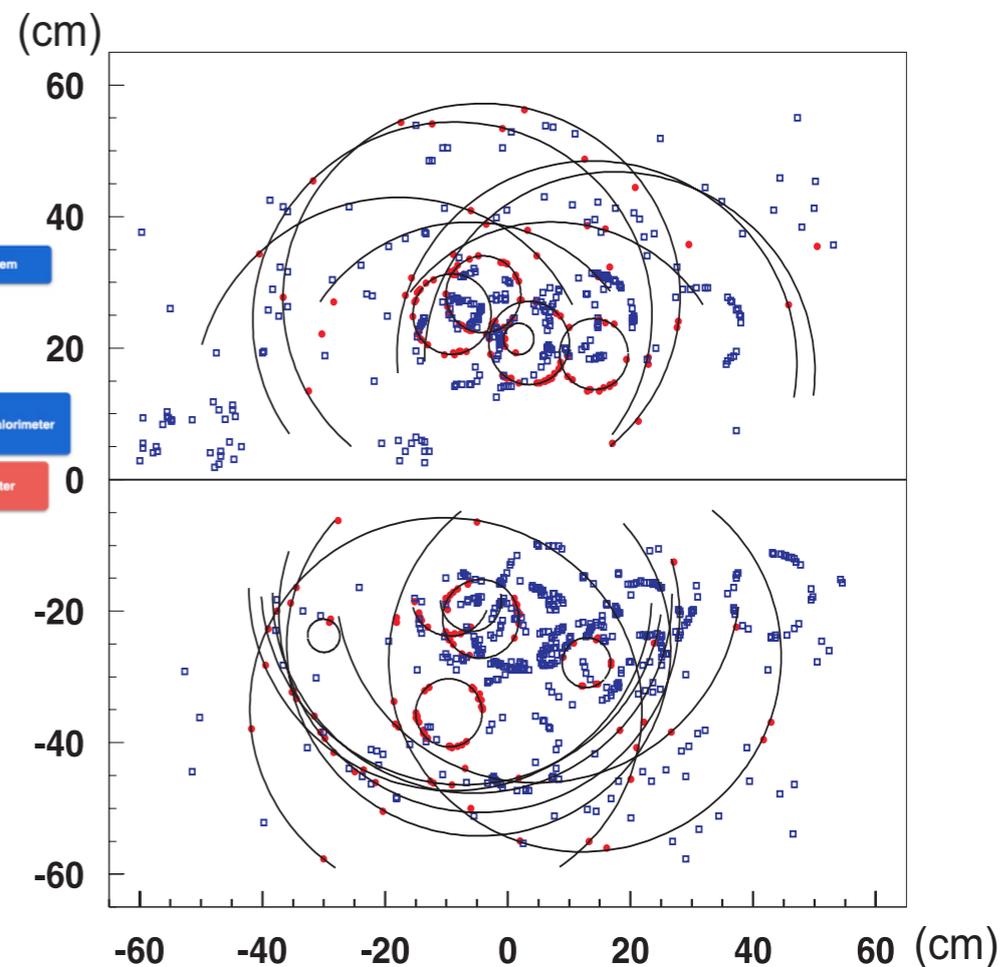
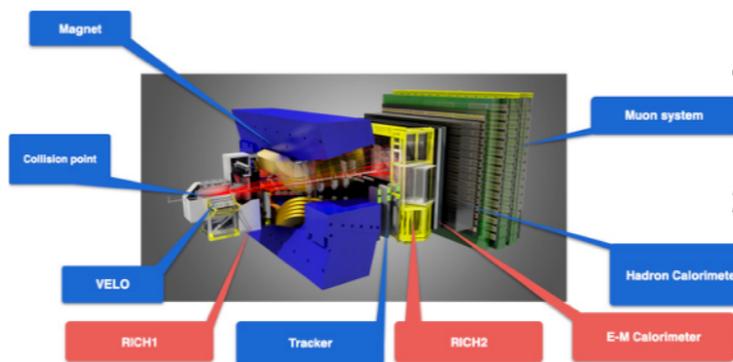
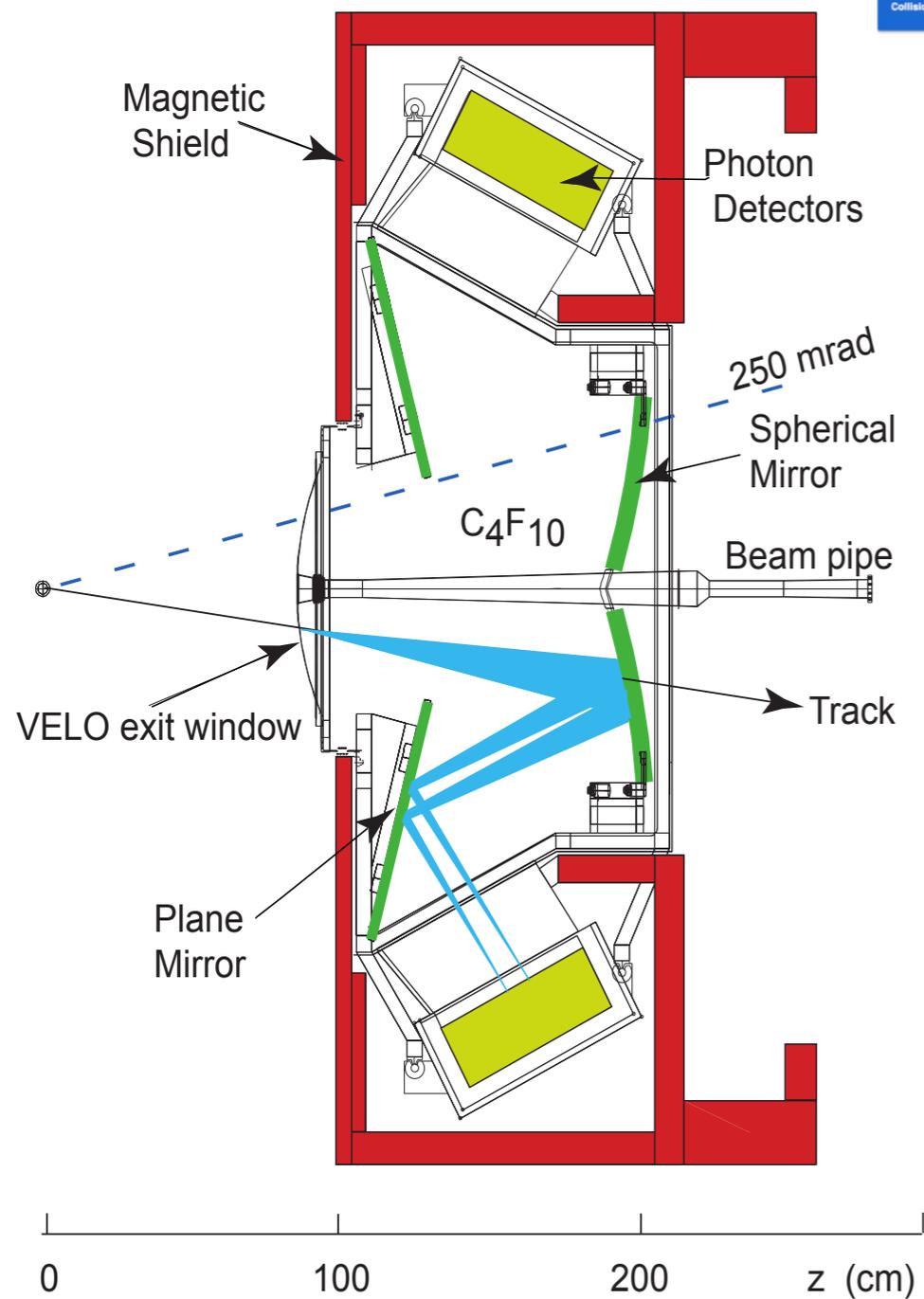
неточности слабого регрессора взаимно сокращаются

Суррогатный регрессор



Приличное качество и при откровенно слабом регрессоре

Детектор RICH



Важнейший прибор для определения типа частиц

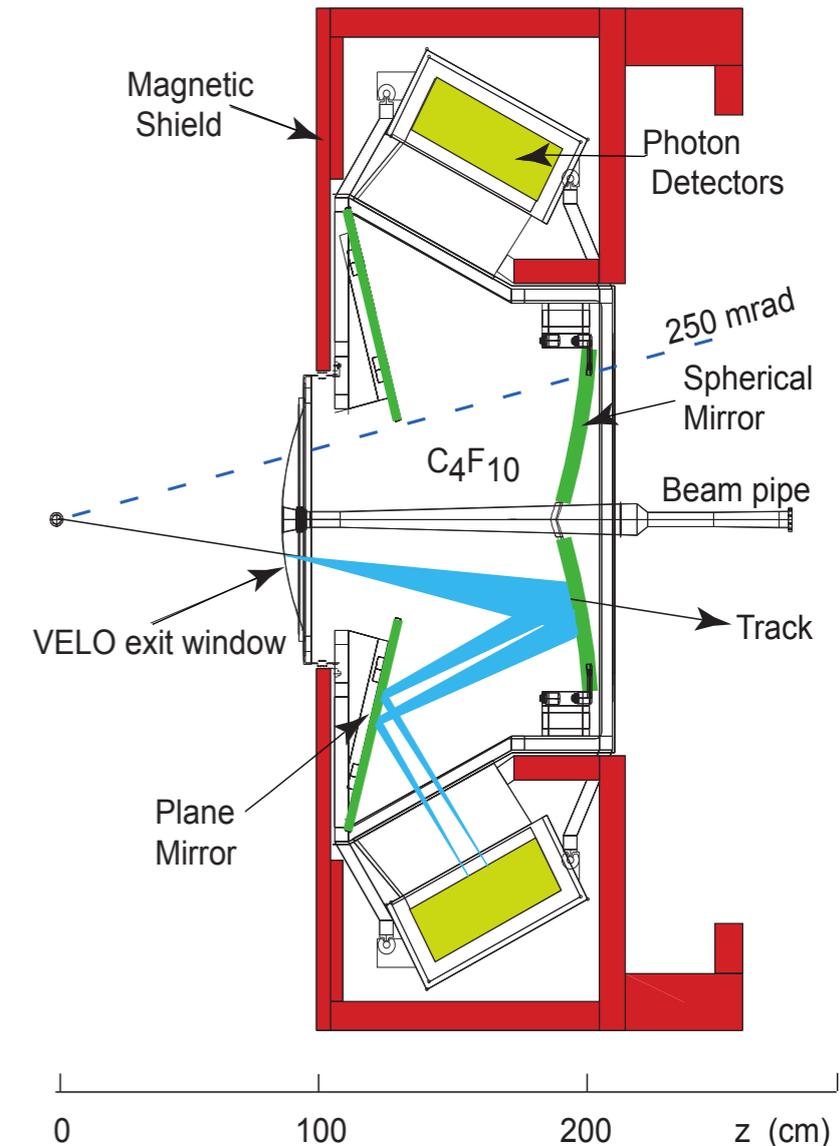
Генеративные модели реальных данных (RICH LHCb)

Симуляция RICH включает:

- симуляцию прохождения частицы через радиатор
- генерацию черенковских фотонов
- распространение фотонов, отражение, переломление, рассеяние
- симуляцию HPD

Все эти вычисления требуют существенных компьютерных ресурсов

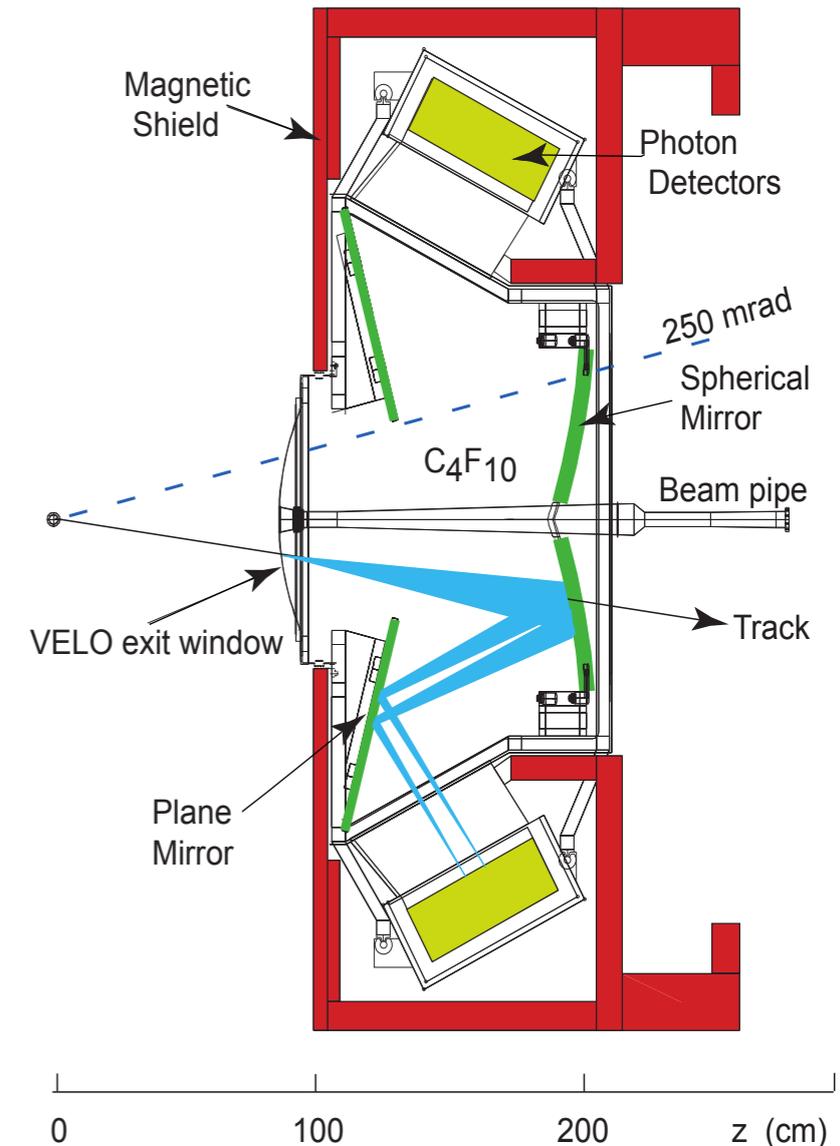
- при этом воспроизводимость получаемых параметров идентификации частиц оказывается не настолько хорошей, как хотелось бы



Генеративные модели реальных данных (RICH LHCb)

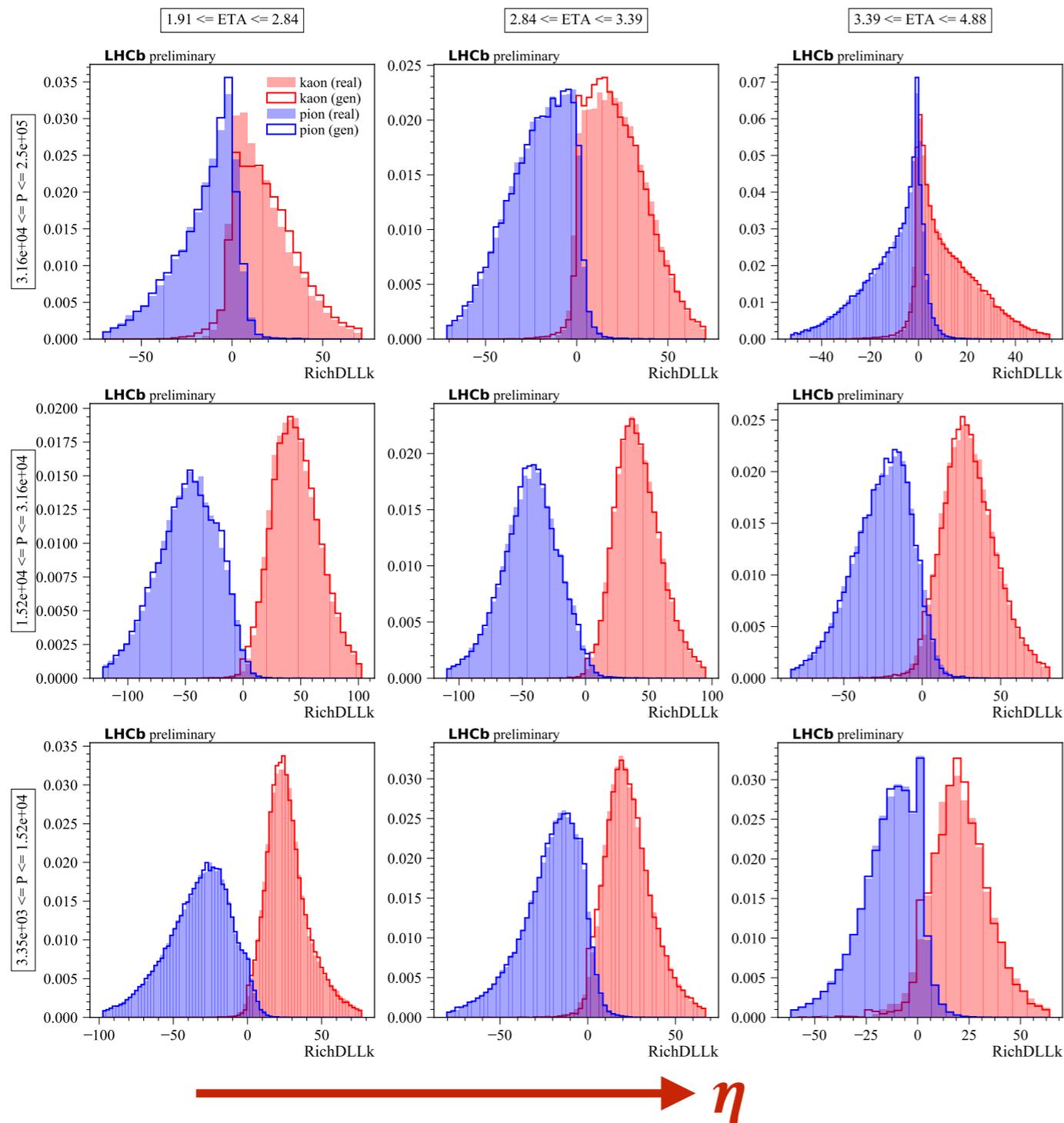
Натренируем генеративную модель переводить кинематику трека непосредственно в параметры идентификации

можно тренироваться прямо на калибровочных образцах реальных (не симулированных) данных



Генеративные модели RICH

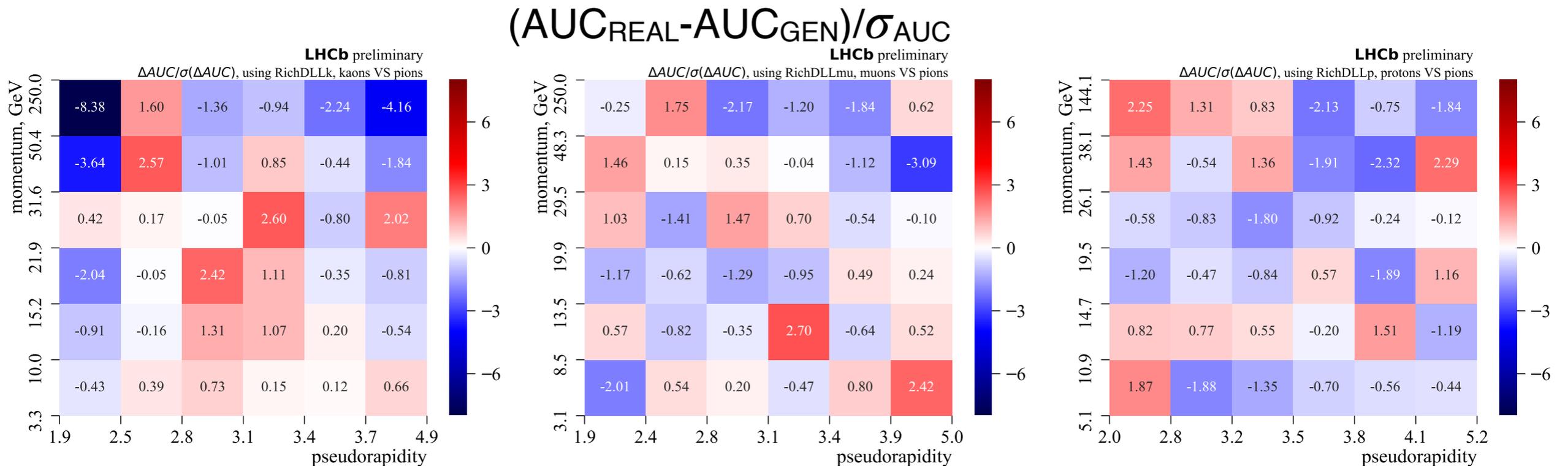
arXiv:1903.11788



Хорошо или плохо?

зависит от практической задачи

Качество генеративной модели



K vs π , using RichDLLk

μ vs π , using RichDLLmu

p vs π , using RichDLLp

Отклик RICH используется для последующей идентификации типов частиц

метрика качества - воспроизведение меры разделяющей способности

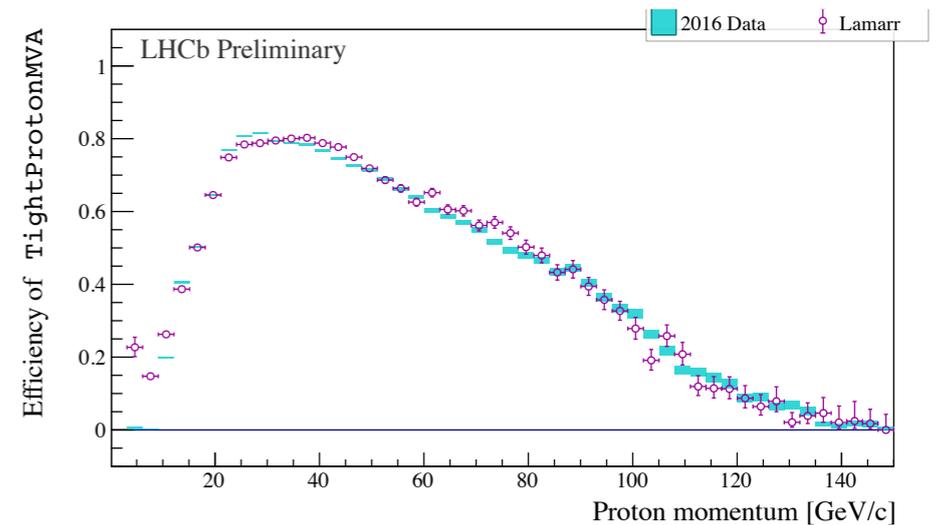
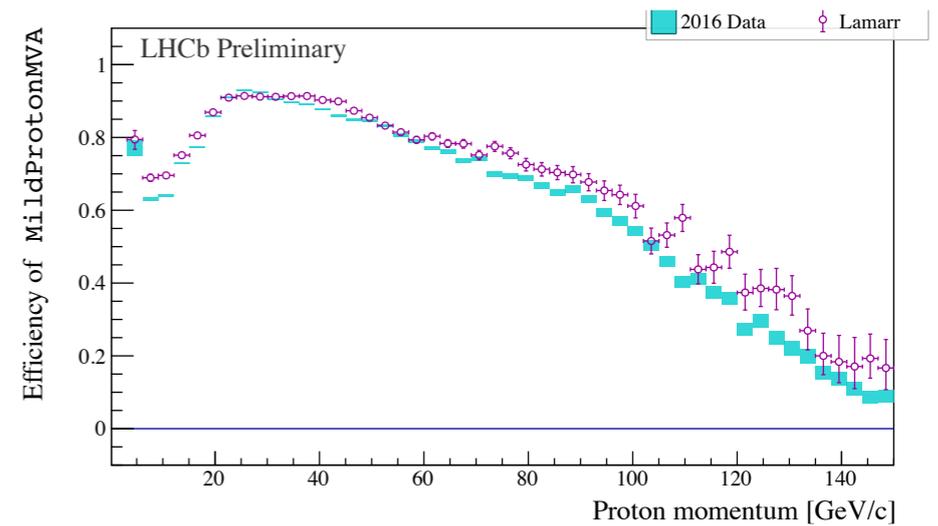
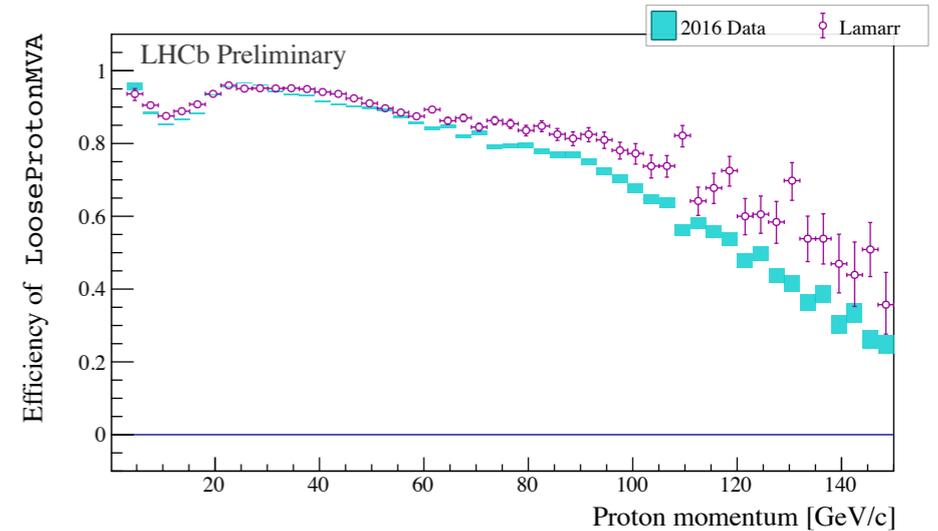
различия учитываются как систематическая ошибка (bias)

получаемого физического результата

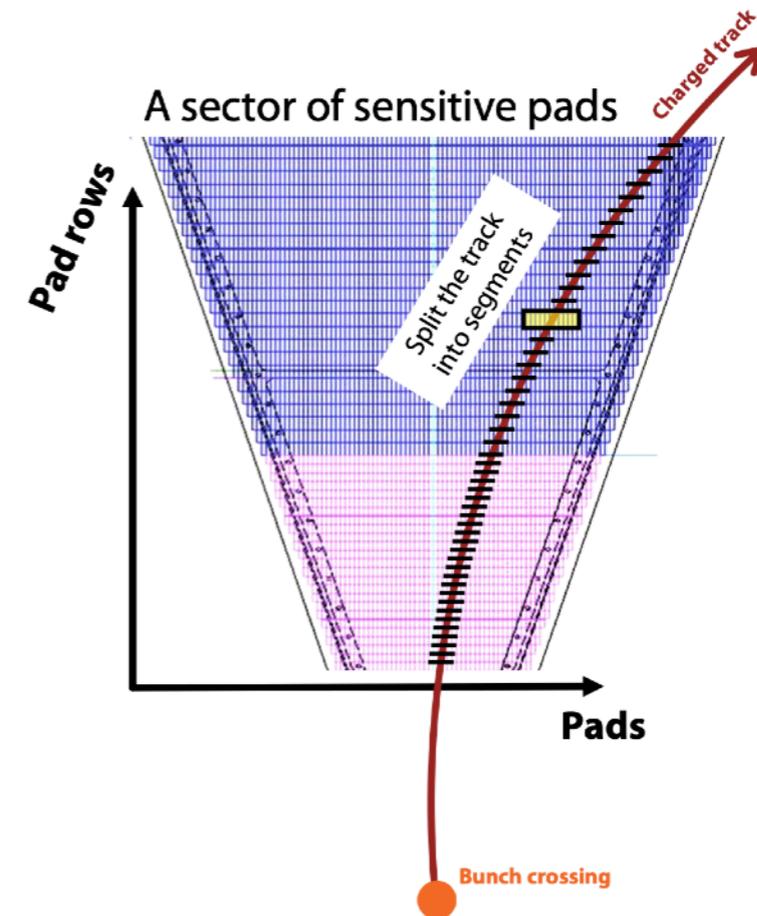
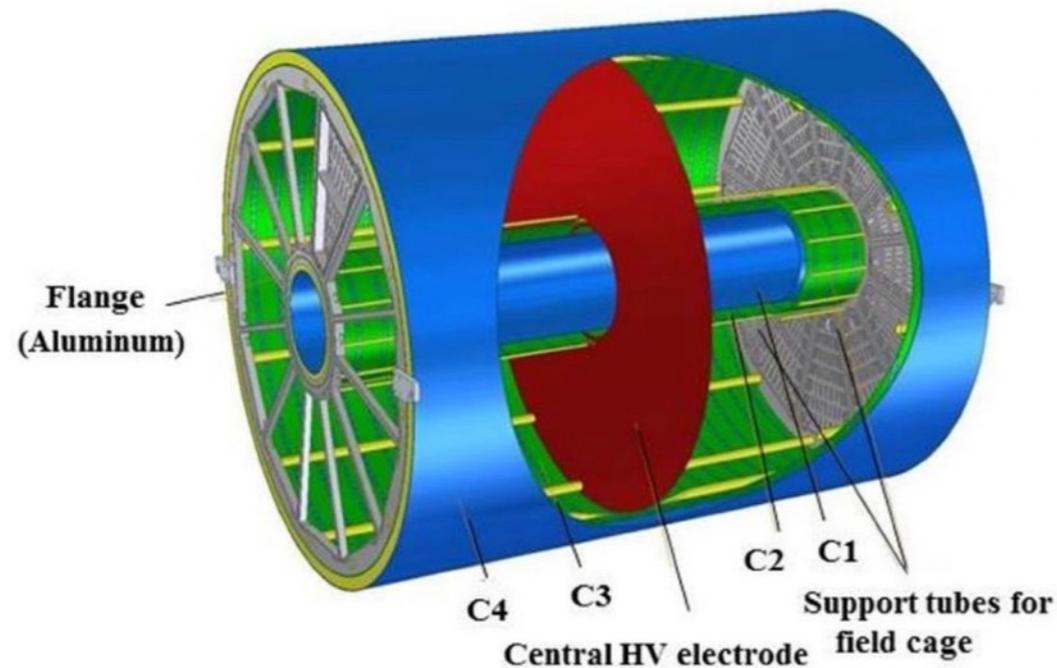
Генеративные модели на реальных данных

NB: сравнение генеративной модели с независимыми данными

малые расхождения уходят в систематику



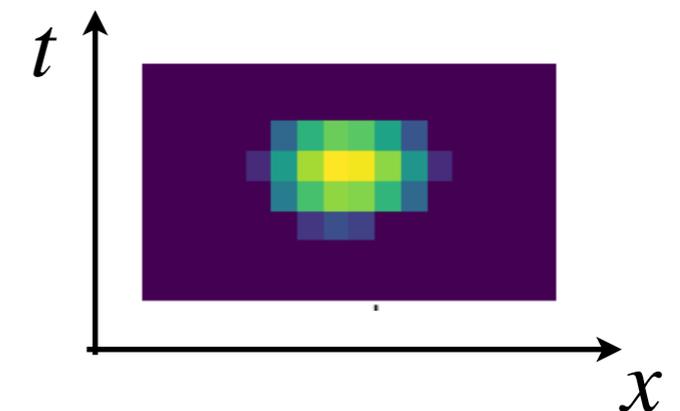
Точность генеративной модели (TPC@MPD)



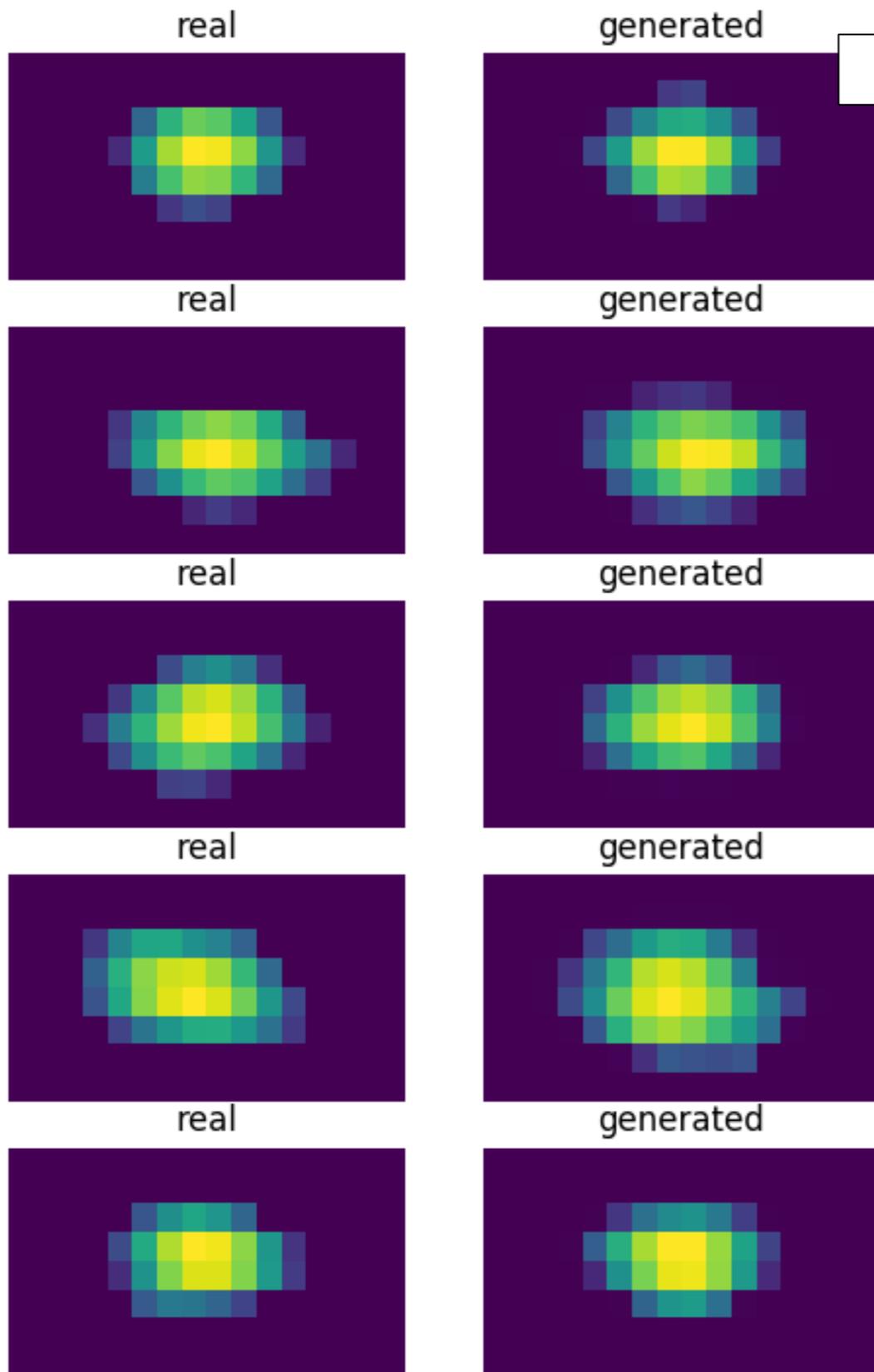
генеративная модель

in: 6 параметров трека
out: стохастический сигнал

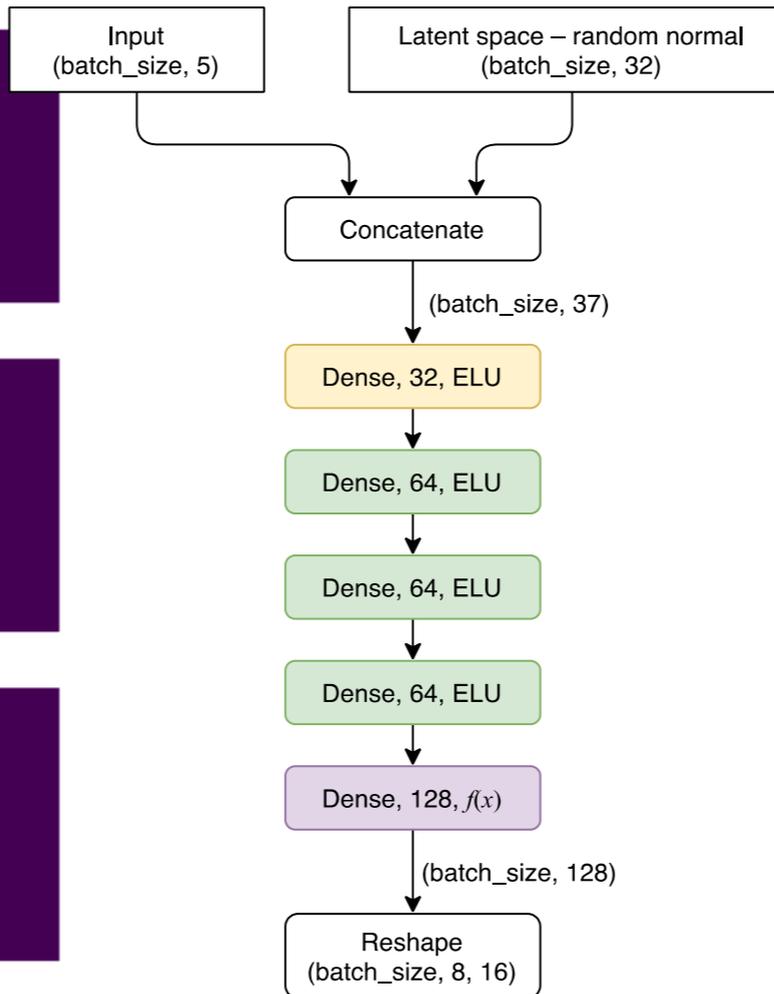
› 8 (пады) × 16 (время) отклик для каждого
ряда падов



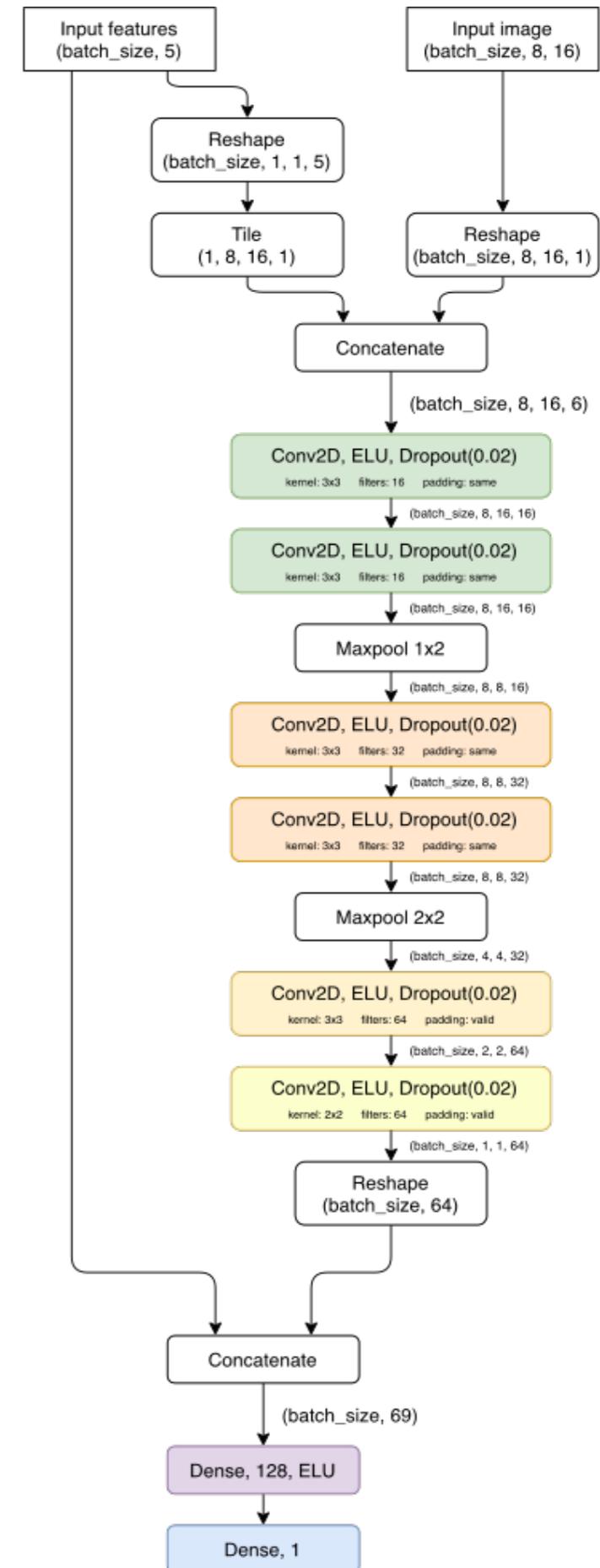
Генератор откликов ТРС



Generator



Discriminator



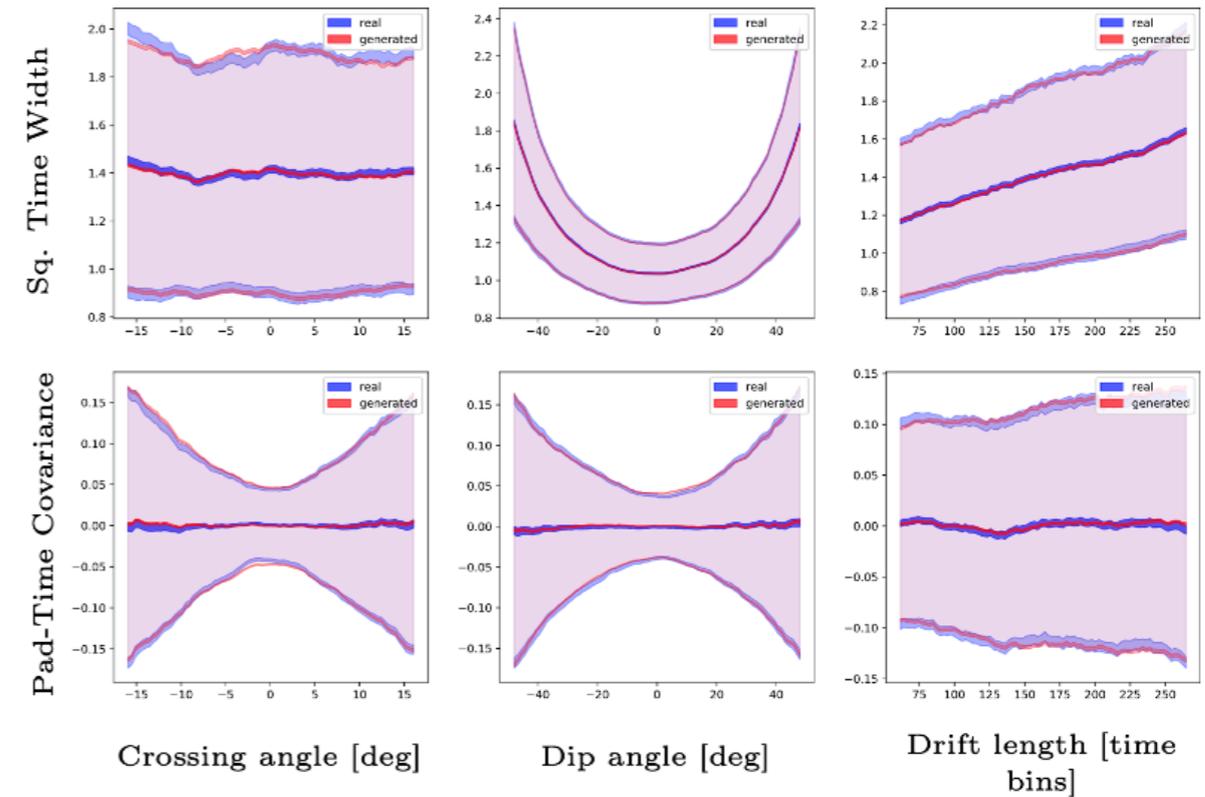
Низко- и высокоуровневая валидация

Eur. Phys. J. C 81, 599 (2021)

Низкий уровень

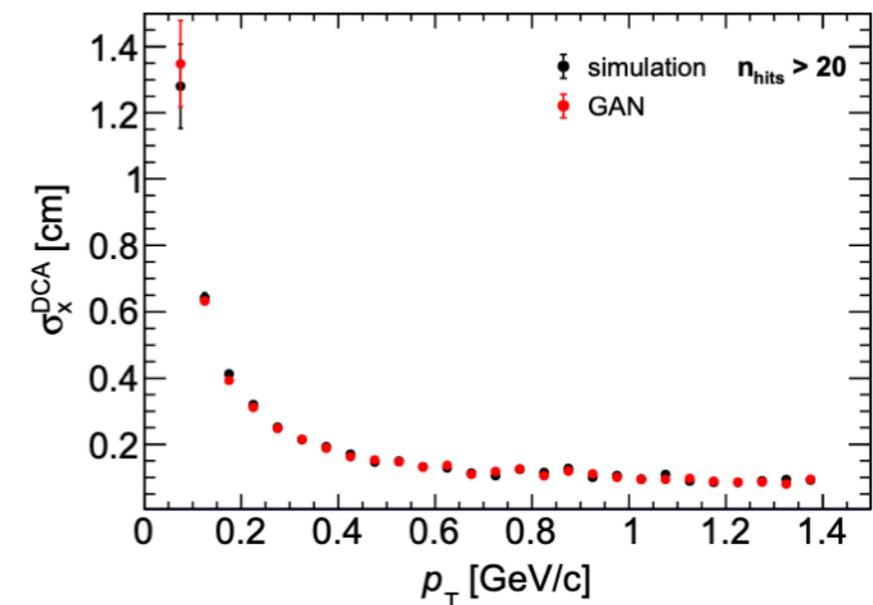
вычисляем первые и вторые моменты сигнального кластера

оцениваем разницу между этим моментами для оригинальных и сгенерированных сигналов для разных значений параметров



Высокий уровень

интегрируем модель в программный стек симуляции детектора
сравниваем качества реконструкции



Distance of closest approach resolution along x

Неопределенности, порождаемые суррогатной моделью

Для естественнонаучного исследования важно не только получить количественный результат, но и оценить интервал его неопределенности

- если при получении результата использовалась суррогатная модель, необходимо оценить возможный bias

Как оценить неопределенность генеративной модели?

Например, ансамблевый подход

- Используем dropout при обучении

 - используем фиксированные DO маски для создания ансамбля моделей

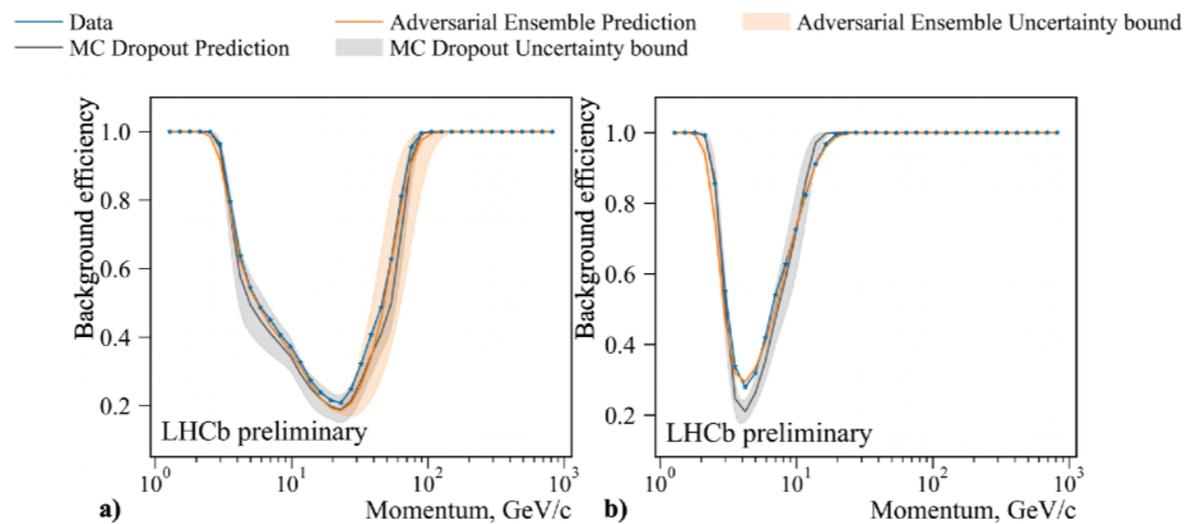
- Штрафуем очередную обучаемую модель за распределения, похожие на ранее обученные модели

 - по мере обучения убираем штраф

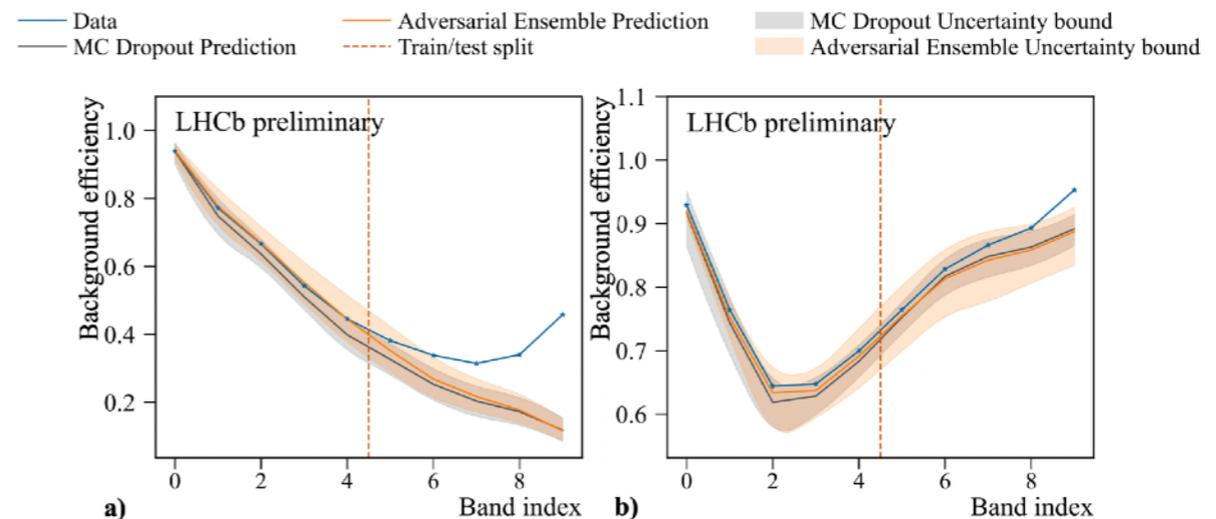
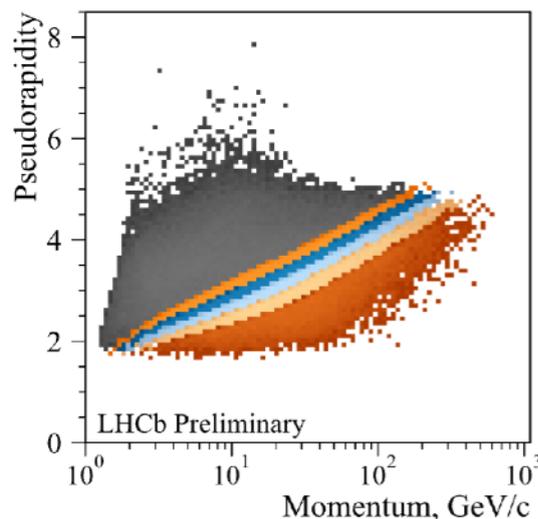
принудительно раздвигаем решения в различные локальные минимумы

Ансамбль обученных моделей позволяет оценить систематические неопределенности, привносимые использованием модели

интерполяция работает разумно



экстраполяция вне домена обучения ненадежна



Заключение

Использование суррогатных генеративных моделей может существенно помочь в научных исследованиях, требующих массовой симуляции данных для интерпретации измерений

Требования к таким моделям сильно отличаются от обычно используемых

- адекватная информативность и вариативность генерируемых объектов

- приоритизация фокуса обучения на физически мотивированные, часто маргинальные метрики

- количественная оценка систематических неопределенностей физических результатов, вносимых вариативностью и смещением модели