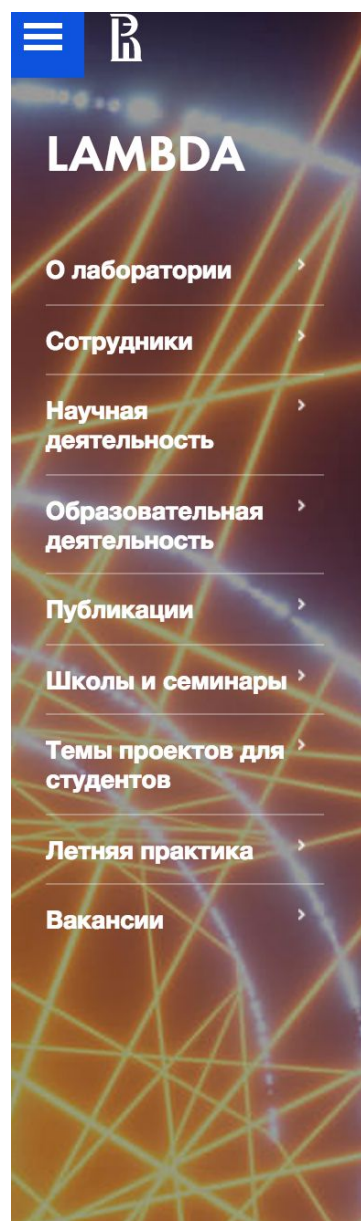


# Network для быстрой генерации откликов калориметров детекторов элементарных частиц в ЦЕРН



Национальный исследовательский университет «Высшая школа экономики» → Факультет компьютерных наук → Научно-учебная лаборатория методов анализа больших данных

RU EN   

## Научно-учебная лаборатория методов анализа больших данных

Принципиальная проблема обработки больших данных заключается в их разнородности и избыточности. Анализ разнородной информации позволяет строить более полное описание окружающей реальности и получать более точные прогнозы. В то же время объемы этих данных растут экспоненциально и требуют специальной подготовки для эффективной работы с ними. С целью создания практического центра международного уровня для повышения мастерства обработки и анализа больших данных в феврале 2015 года была организована Лаборатория Методов анализа Больших Данных (LAMBDA).

В настоящее время лаборатория фокусируется на применении методов машинного обучения и анализа данных для решения задач фундаментальных наук, таких как физика частиц и астрофизика. Совместный поиск ответов на загадки Вселенной с ведущими учеными из этих областей составляет основное направление развития лаборатории. В частности, лаборатория сотрудничает с Европейским центром ядерных исследований (CERN), и совместная исследовательская работа заключается как в исследованиях физики событий Большого адронного

[Подписаться на анонсы](#)

### КОНТАКТЫ



[Устюжанин Андрей Евгеньевич](#) —  
заведующий лабораторией  
[austyuzhanin@hse.ru](mailto:austyuzhanin@hse.ru)



[Глазистов Артём Владимирович](#) —  
менеджер лаборатории  
[aglazistov@hse.ru](mailto:aglazistov@hse.ru)



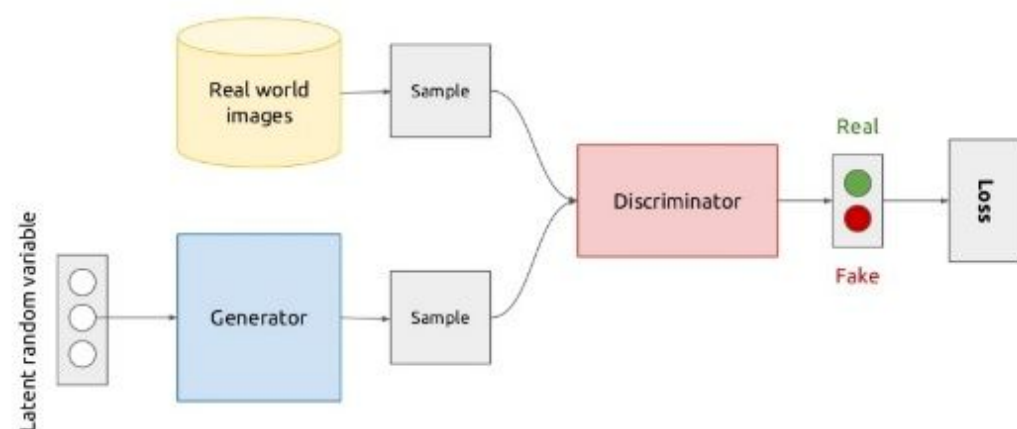
# GAN

## ◇ Generative Adversarial Networks

◇ современная технология использования машинного обучения без учителя

◇ генерация объектов с заданными свойствами

### Adversarial Learning



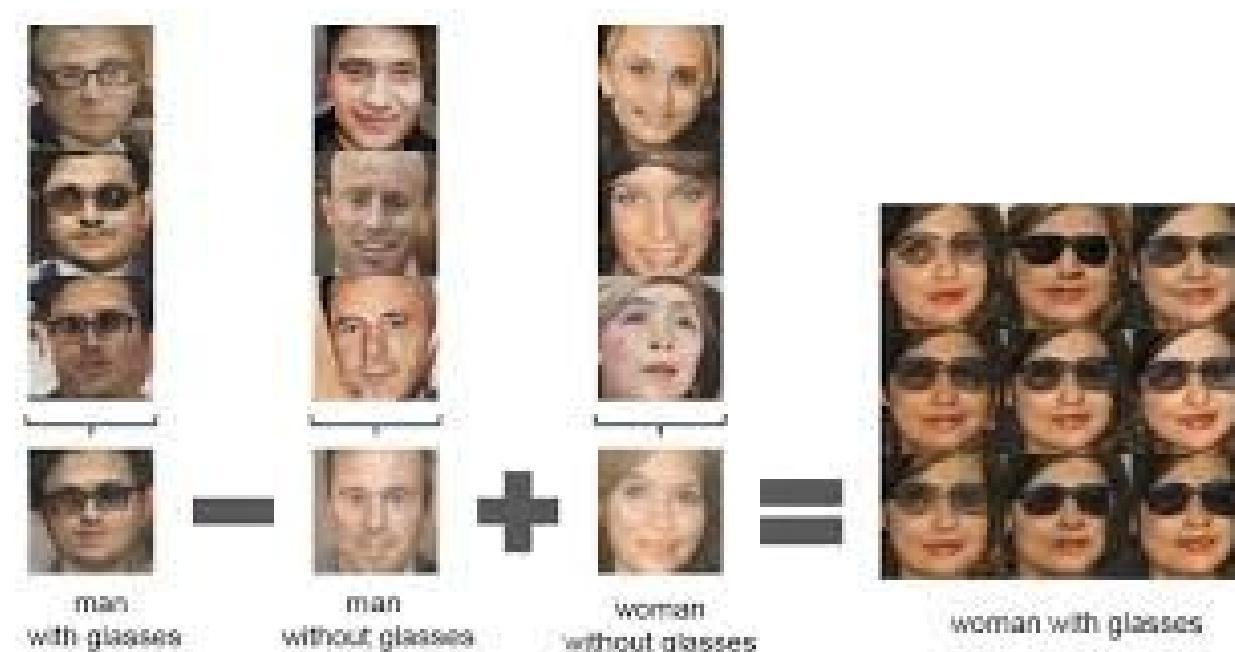
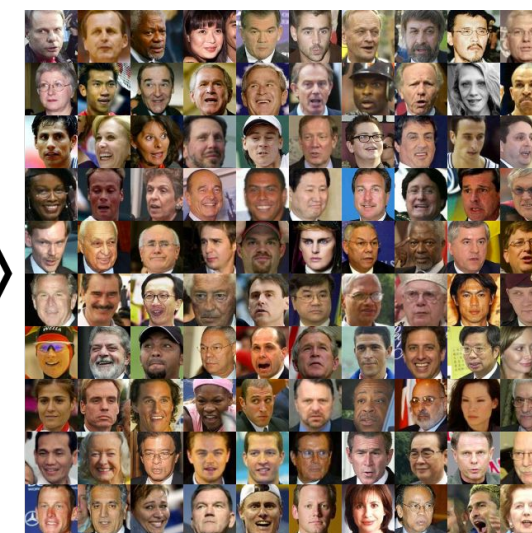
<http://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016>



Noise  $\sim N(0,1)$



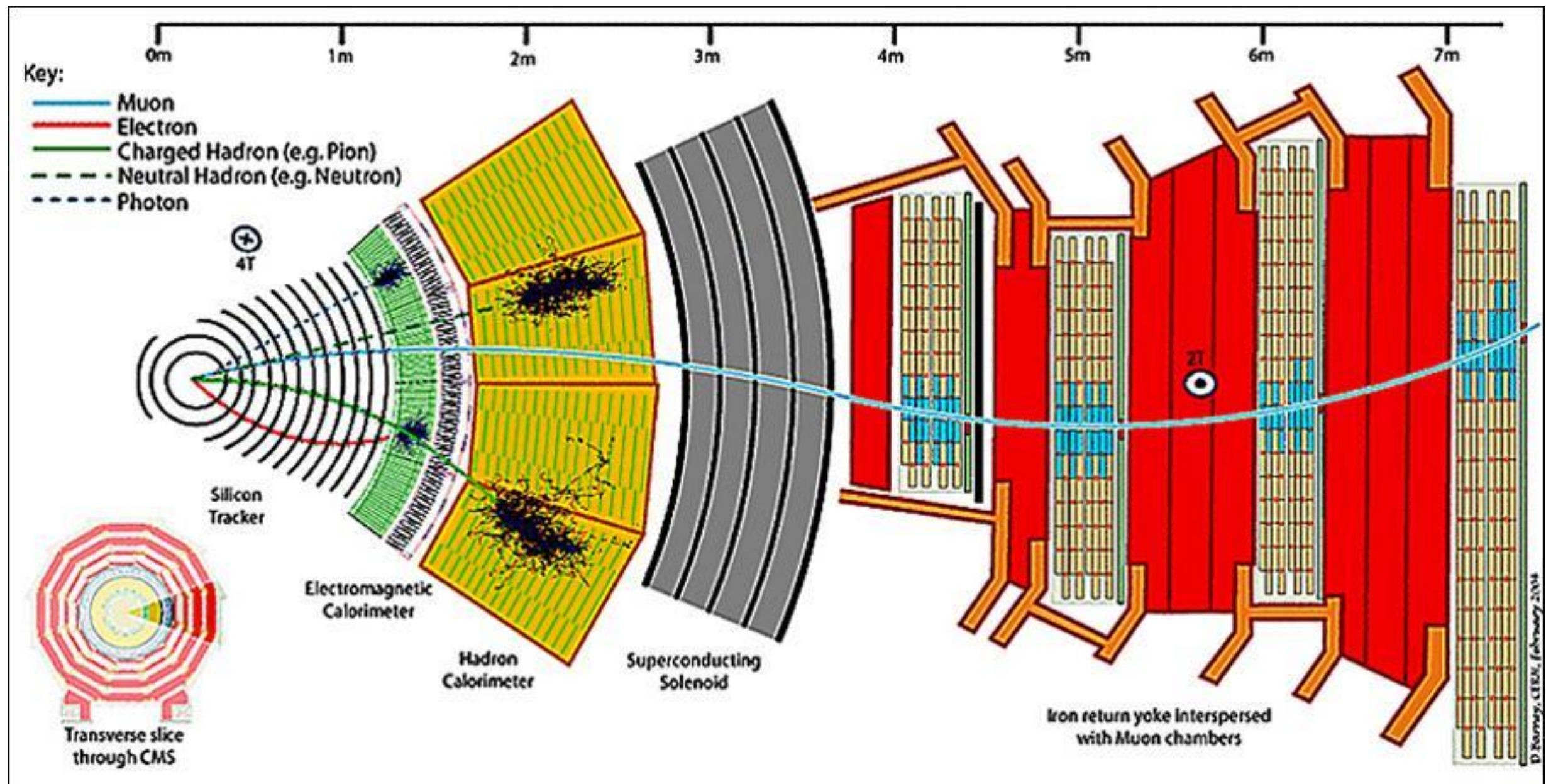
Generative Model



- ◇ 2 протонных пучка, скорость - 99.9999991% от скорости света
- ◇ длина кольца - 27 км, сверхпроводящие магниты
- ◇ 4 точки пересечения пучков
- ◇ энергопотребление max 180 МВт (летом)
- ◇ энергия в пучках ~ энергии Титаника на полном ходу

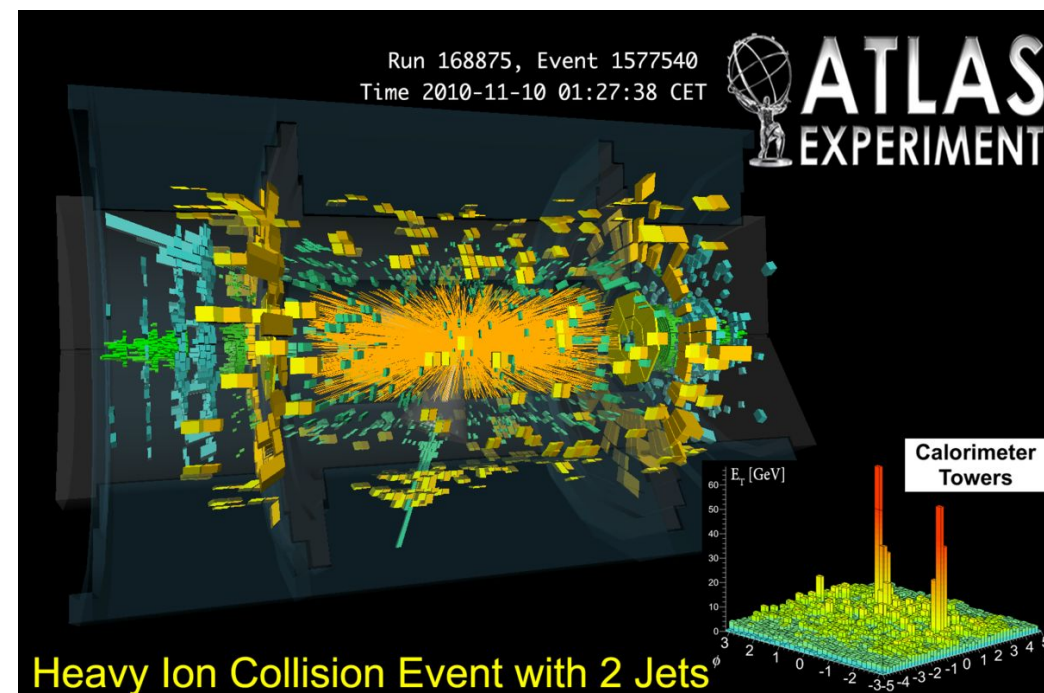


# Поведение частиц в детекторах





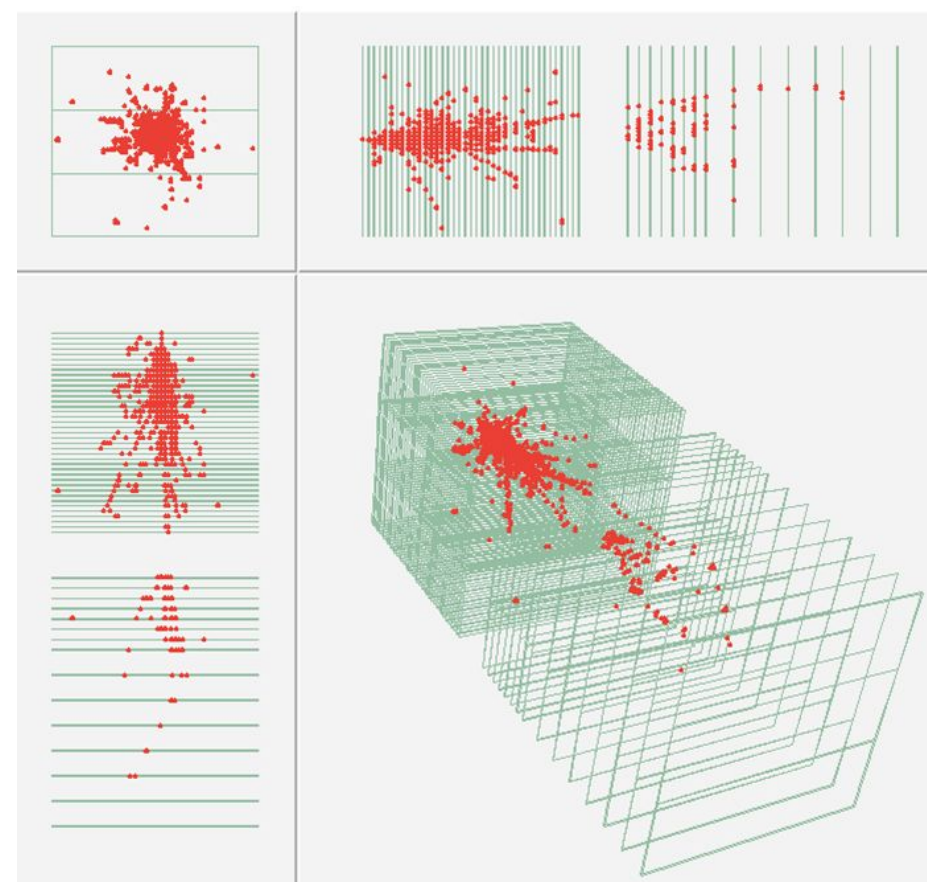
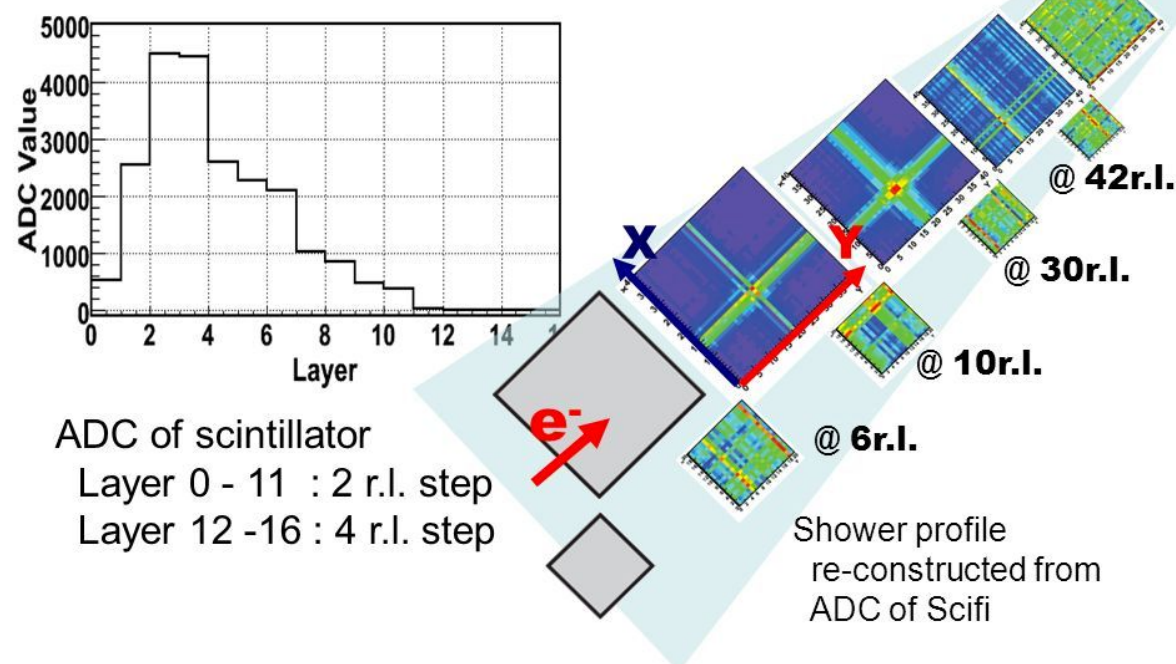
# Калориметры



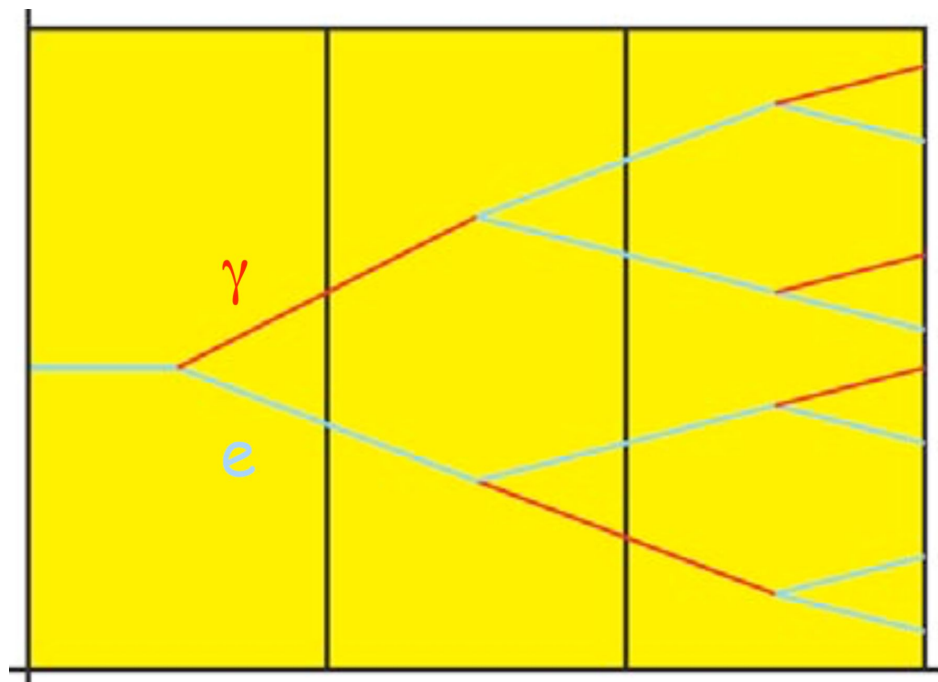
## Event View - electron -

LHCf

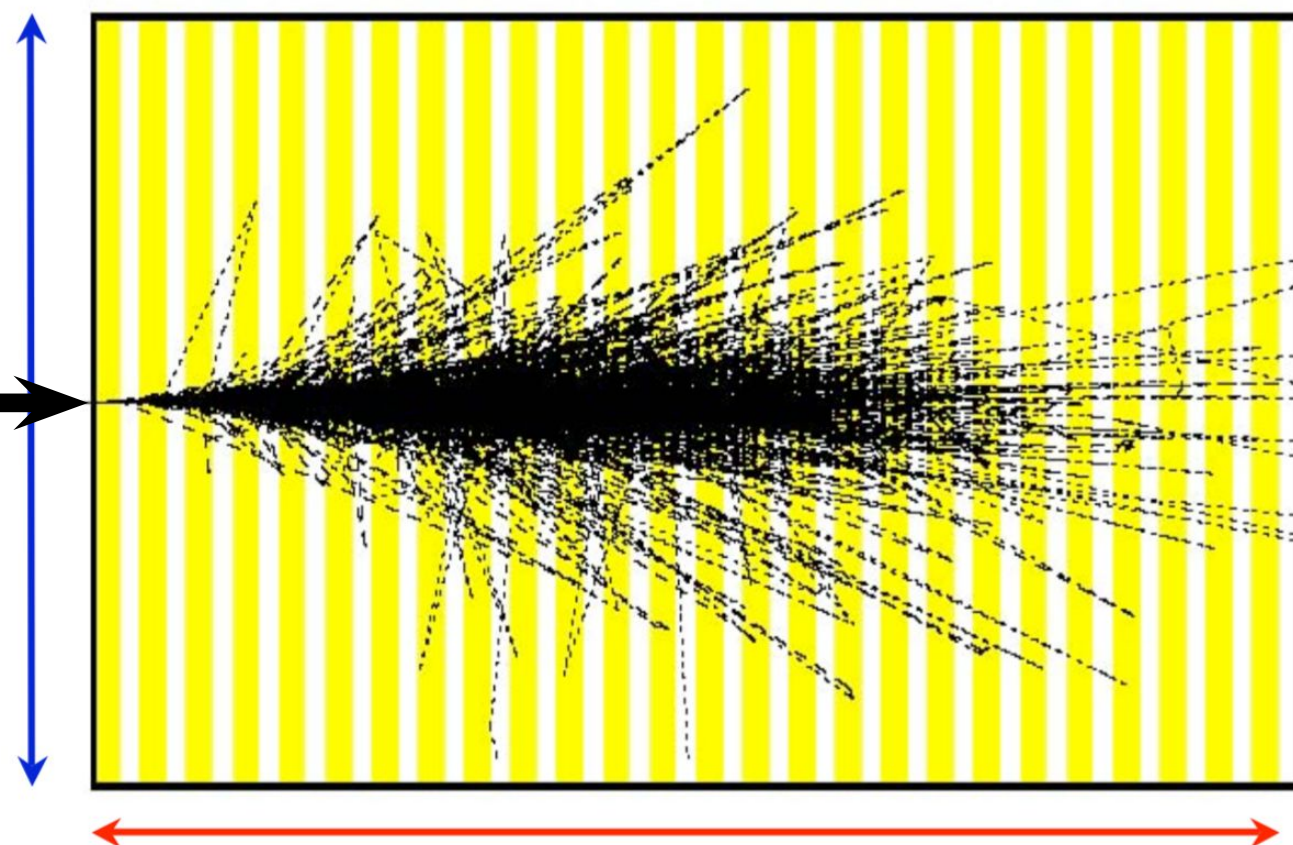
### 200GeV/c Electron Beam Event @ 40mm Calorimeter.



# Ливни в калориметрах



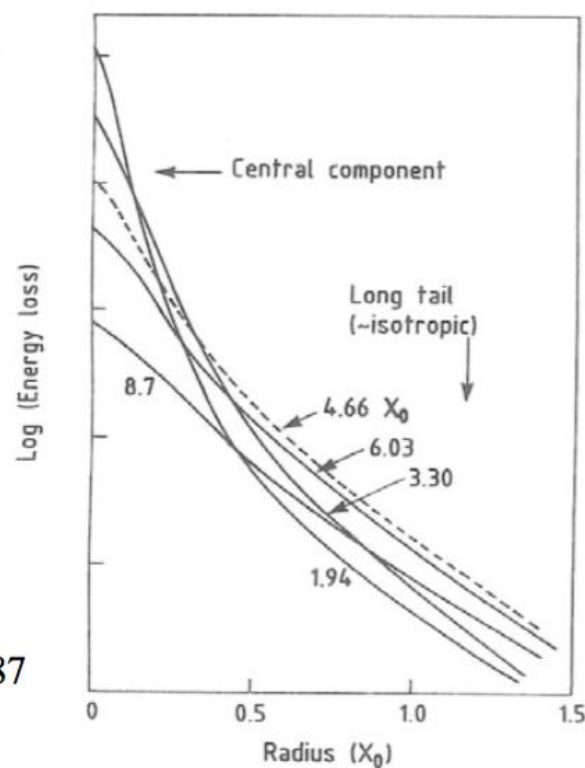
electron,  $E=27.5 \text{ GeV}/c^2$





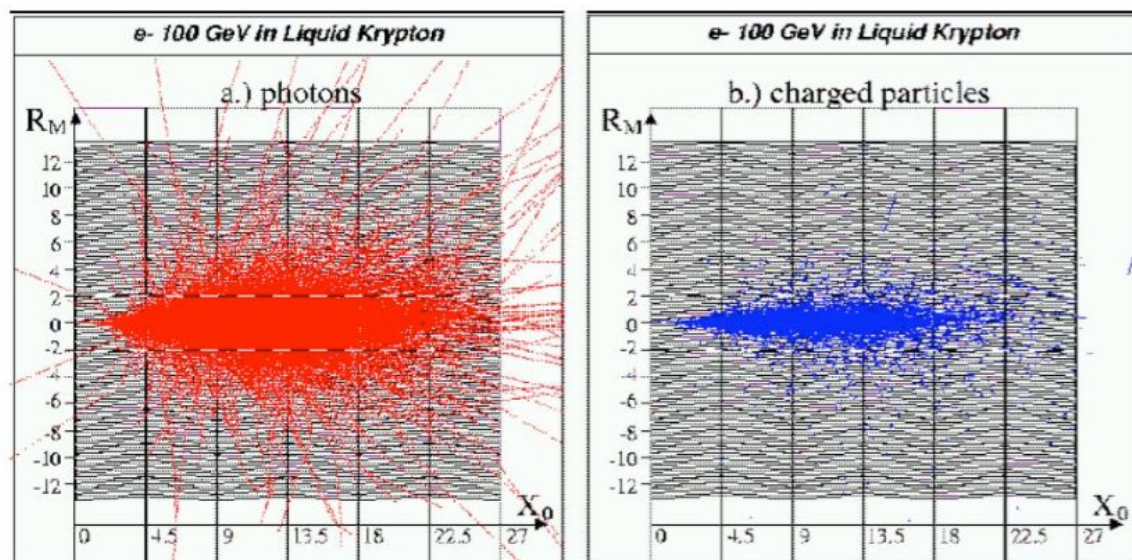
# Electromagnetic showers

## Shower profile

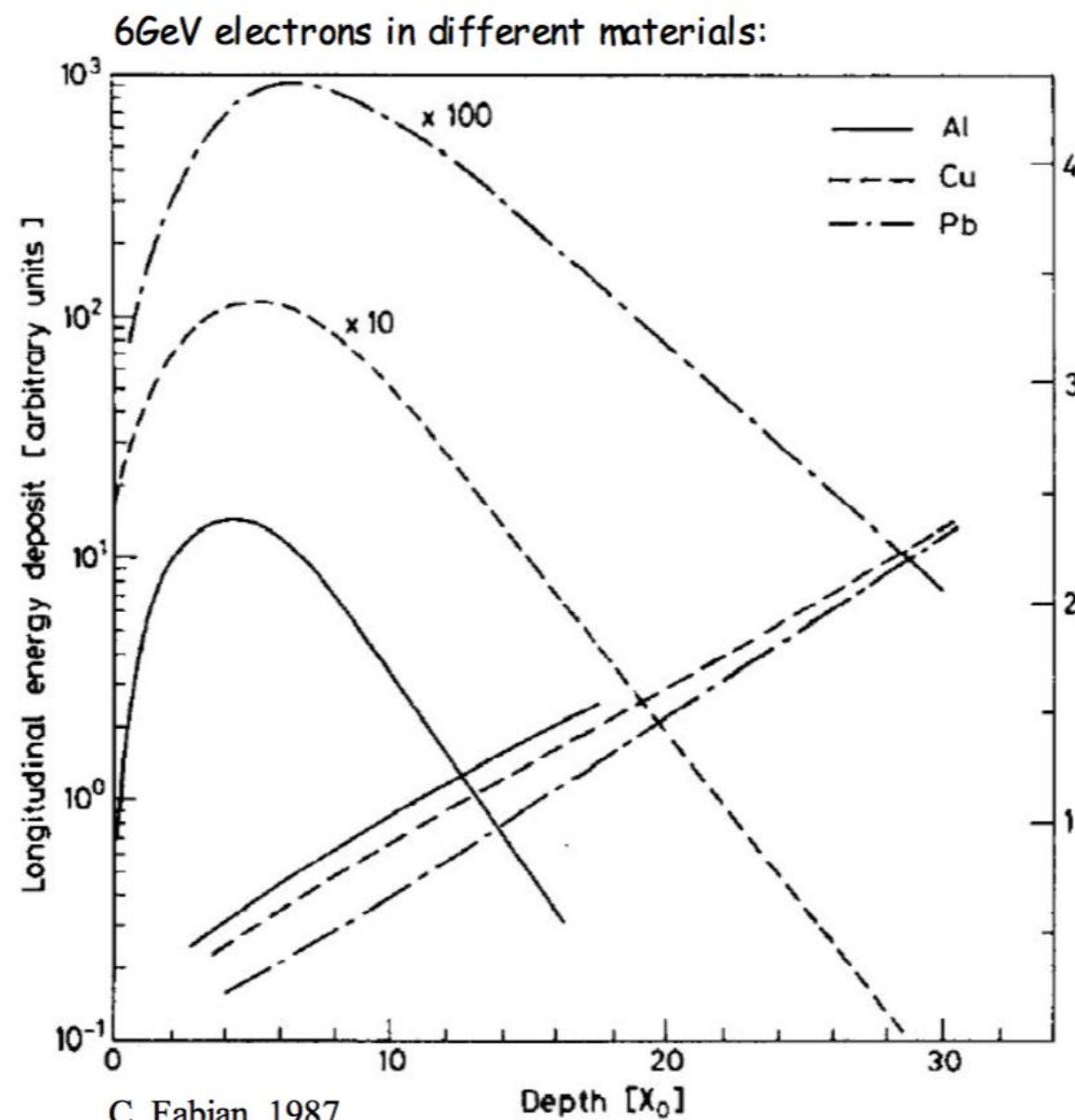


C. Fabjan, 1987

## MC simulation:

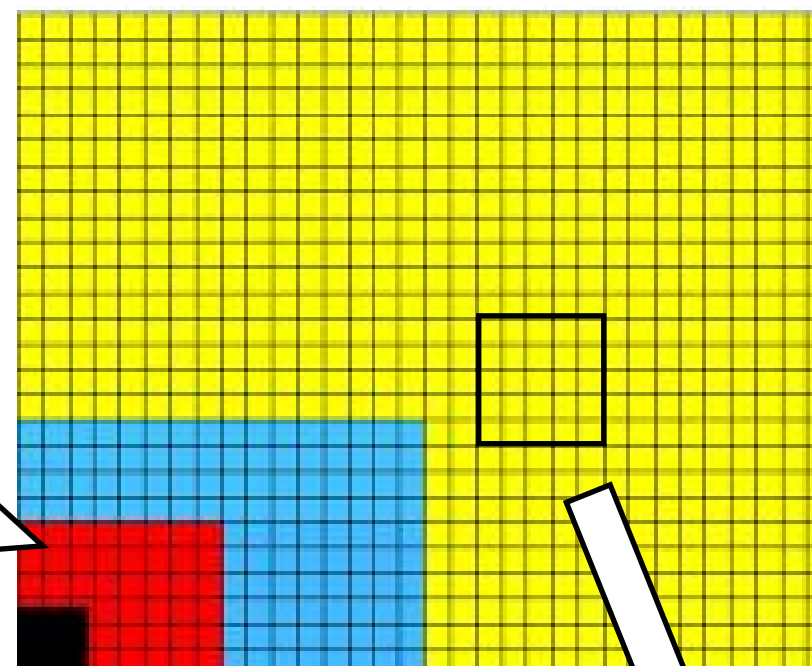


D. Wegener, 2001



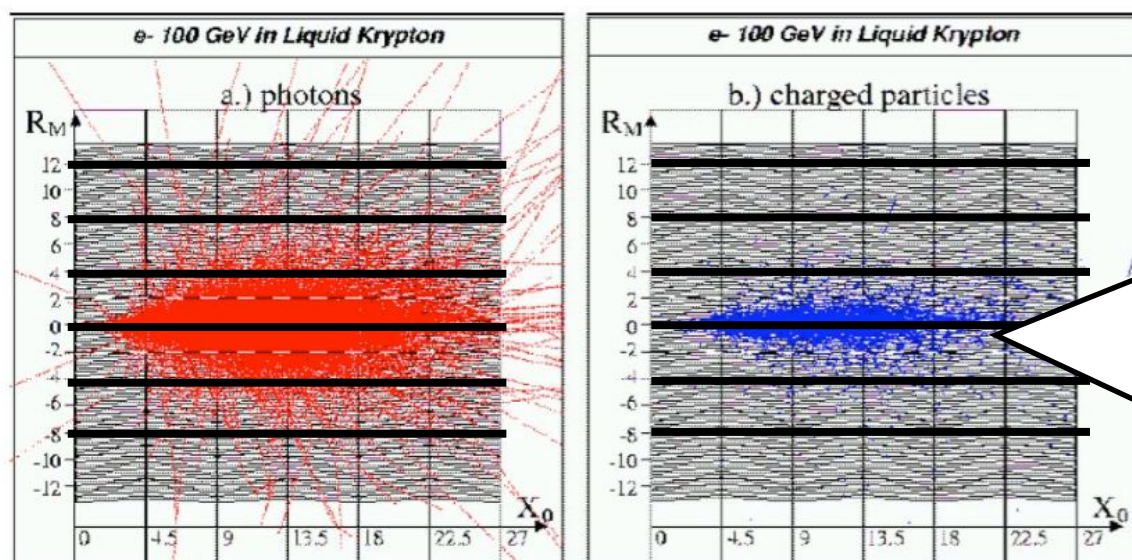
C. Fabjan, 1987

Bernd Su

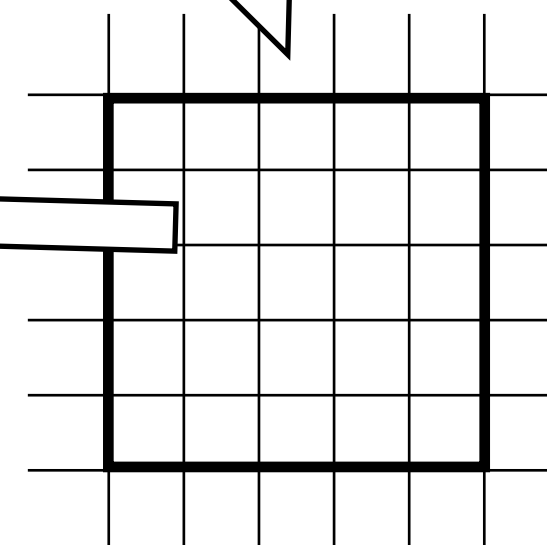


Outer section :
121.2 mm cells
2688 channels
Middle section :
60.6 mm cells
1792 channels
Inner section :
40.4 mm cells
1472 channels

MC simulation:

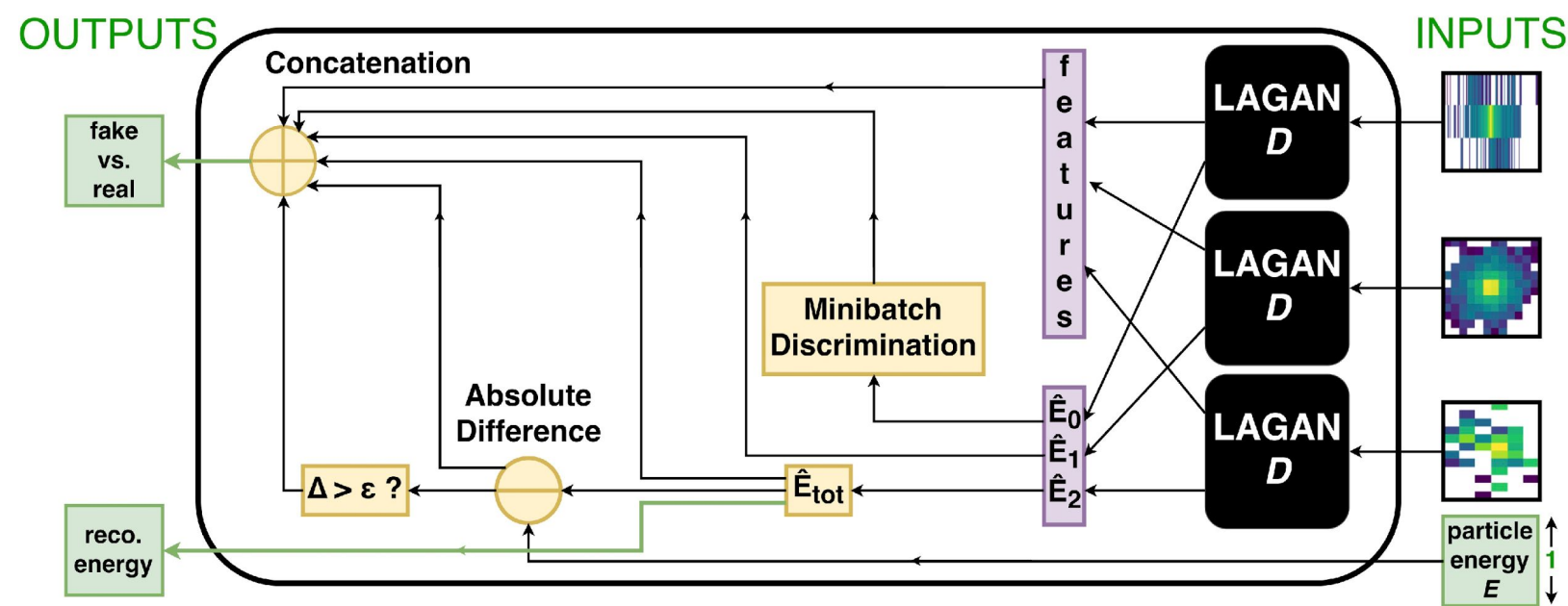
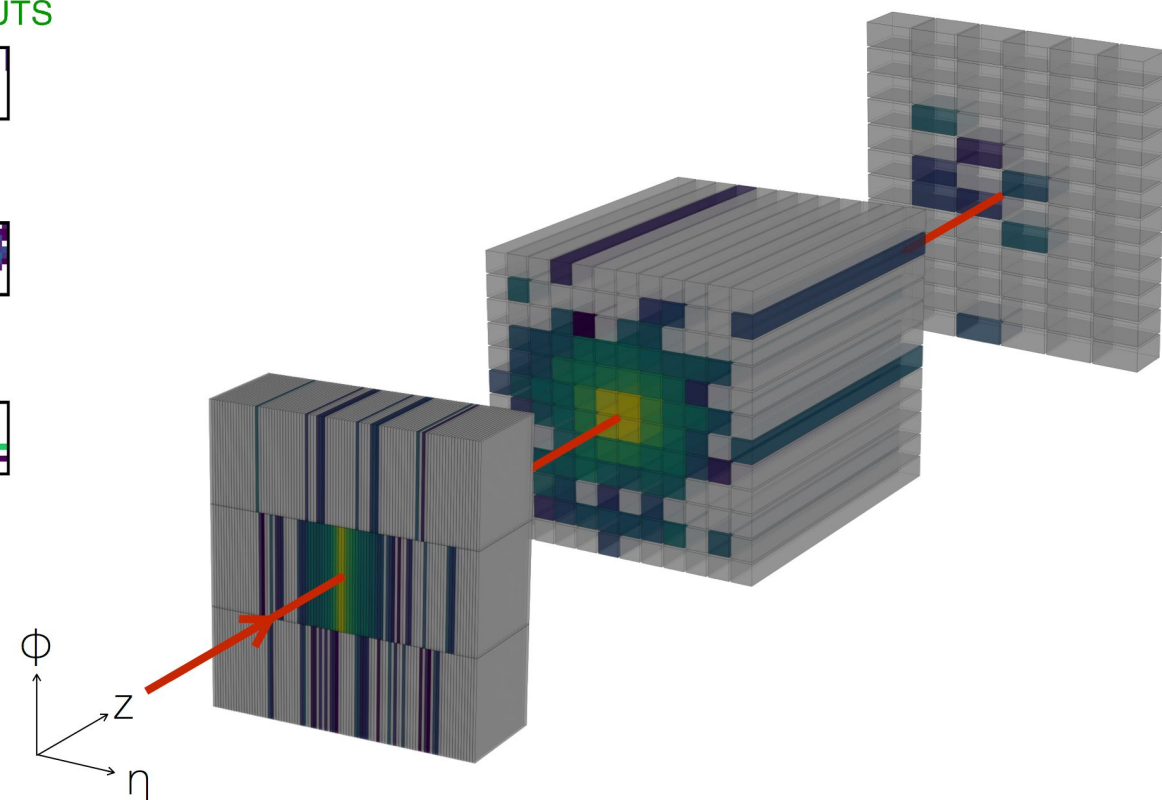
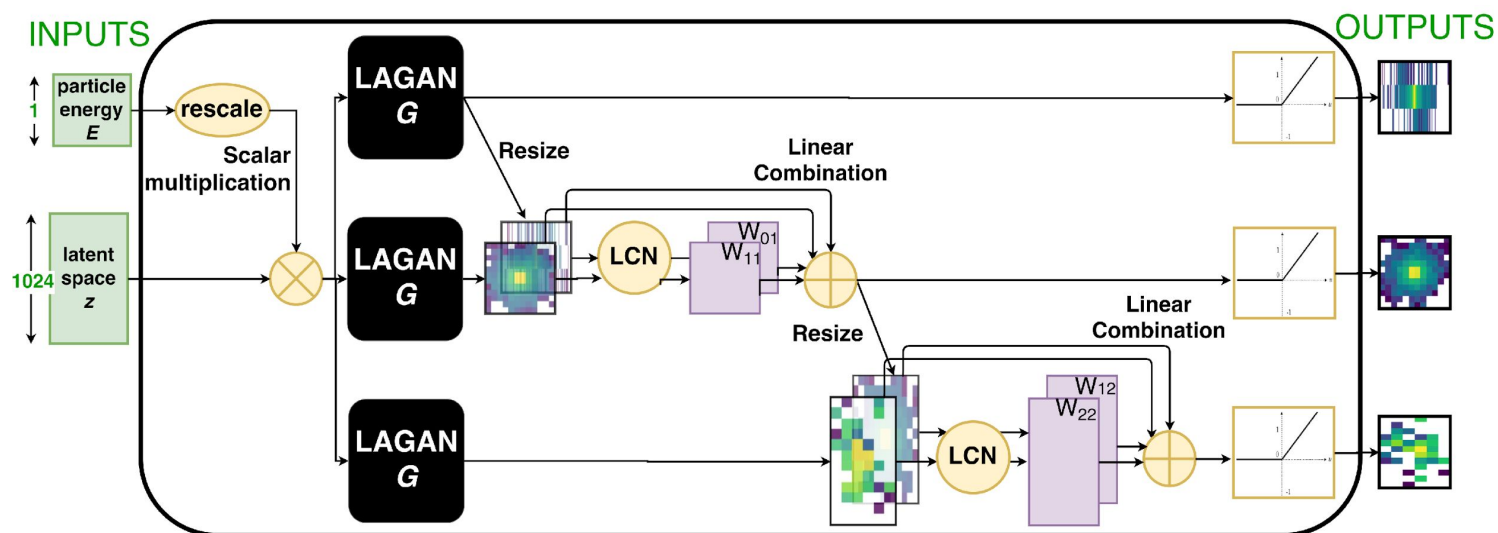


D. Wegener, 2001





# CaloGAN



# Проект

- ◆ Конкретная имплементация калориметра, отклик на конкретный тип частиц
  - ◆ результат полной MC симуляции
- ◆ Фичи от 1 (энергия) до 5 (+2 координаты +2 угла)
- ◆ Построить GAN
- ◆ Обучить GAN
- ◆ Проверить качество генератора
- ◆ На 5+
  - ◆ найти “разумную” метрику успеха
  - ◆ измерить “качество” количественно
  - ◆ оценить систематические погрешности



# Детали

- ◆ Чему научатся студенты? Что самое интересное в проекте?
  - ◆ Проект даёт возможность познакомиться с основами обучения без учителя и разобраться в концепции Generative adversarial networks, а также разобраться со структурой экспериментов Большого адронного коллайдера
- ◆ Компоненты (Из каких частей состоит проект?)
  - ◆ 1) Воспроизведение бейзлайна на существующих данных
  - ◆ 2) Улучшение бейзлайна
  - ◆ 3) Учет нюансов отдельно взятого детектора
- ◆ Какие начальные требования?
  - ◆ Хорошее знание математического анализа и математической статистики
  - ◆ Хорошее знание школьного курса физики
  - ◆ Умение писать и читать код на python (опционально - использовать jupyter, docker)
  - ◆ Не пугаться Git

- ◇ Поскольку задача актуальна и калориметры реальны, в случае успешного доведения решения до production quality оно будет адаптирован коллаборациями в ЦЕРН с соответствующими кредитами, конференциями, публикациями и т.п.