SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

UNIVERSITÉ DE GENÈVE

ESR 3 : Real Time Analysis Strategies for Reconstruction, Exotic Physics and Market Analysis

Leon Bozianu

Supervisors: Anna Sfyrla, Steven Schramm

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Introduction

A little about me:

- Leon Bozianu, originally from the UK.
- Université de Genève PhD student.
- MSc in Physics from Niels Bohr Institute, Copenhagen (June 2022)
- BSc in Mathematics and Physics from Durham University (July 2020)

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

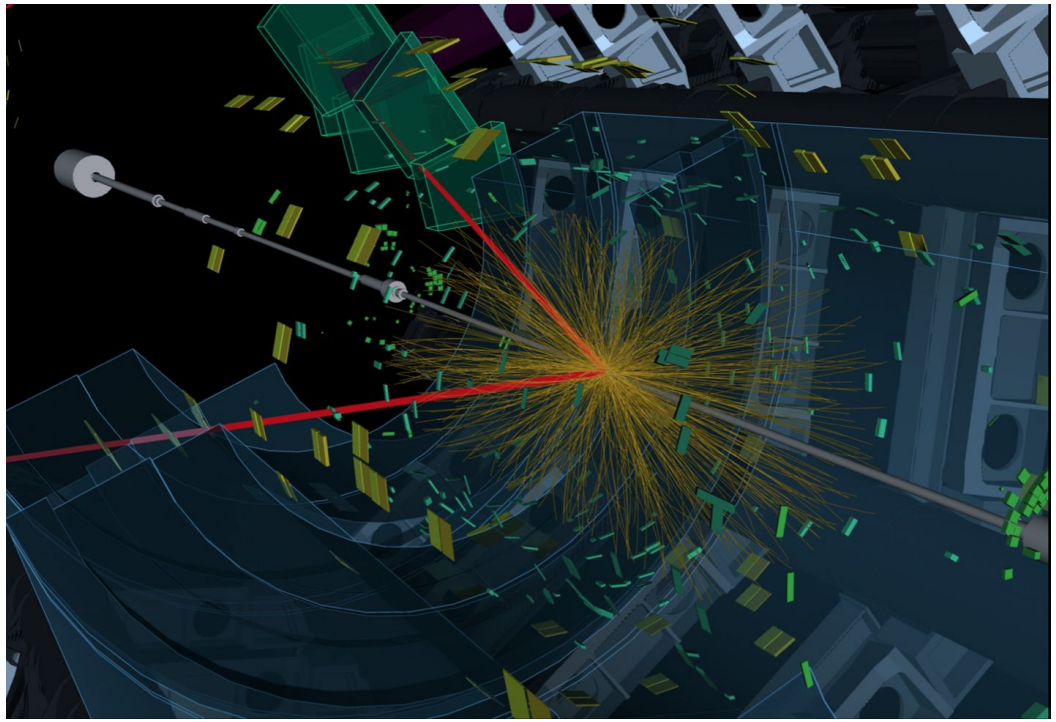
MARIE CURIE ACTIONS

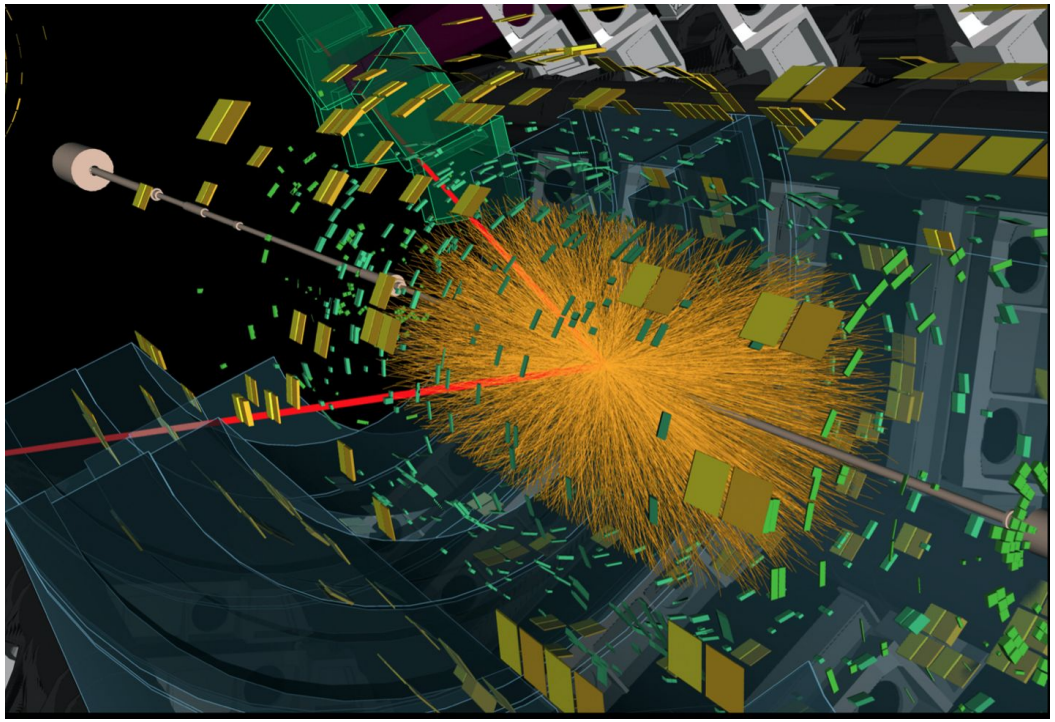
2

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Project Work

We have a pile-up problem!
Currently 50 p-p collisions per event, but forecast to rise to ~200.



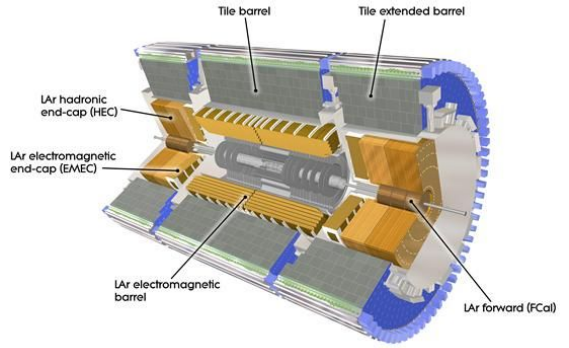


SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

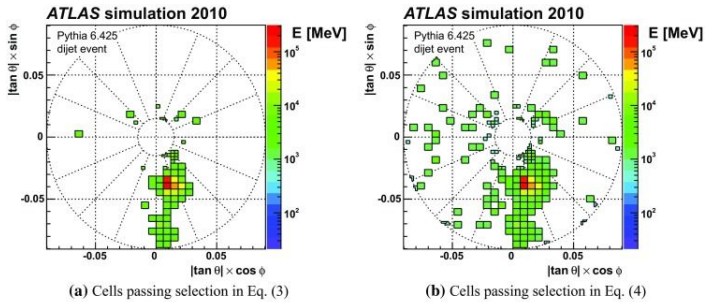
MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Problem Statement

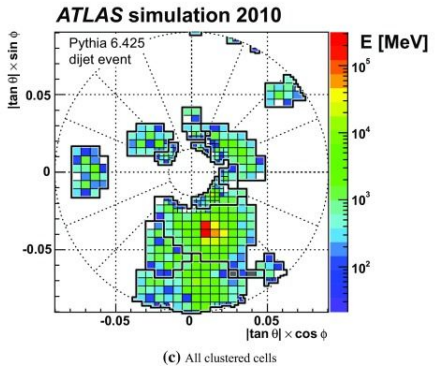


- The TopoClustering algorithm is one of the most resource intensive algorithms used in the high level trigger (HLT).
- The TopoClusters produced in the calorimeter play an important role in jet and MET reconstruction.
- Can we replicate (or improve upon) the results of the current algorithm but faster?



(a) Cells passing selection in Eq. (3)

(b) Cells passing selection in Eq. (4)



(c) All clustered cells

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Project Objectives

- ➔ Investigate ML clustering substitutes for the TopoCluster algorithm.
- ➔ Perform benchmark tests of performance and speed.
- ➔ Potentially extendable to parallel computing structures or GPUs.
- ➔ Additional functionality: cluster classification, energy/momentum estimation.

Any solution must be robust to the granularity and irregularity of the calorimeter design. The necessary training data will be the calorimeter cells from each simulated event (possibility to train in data).

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

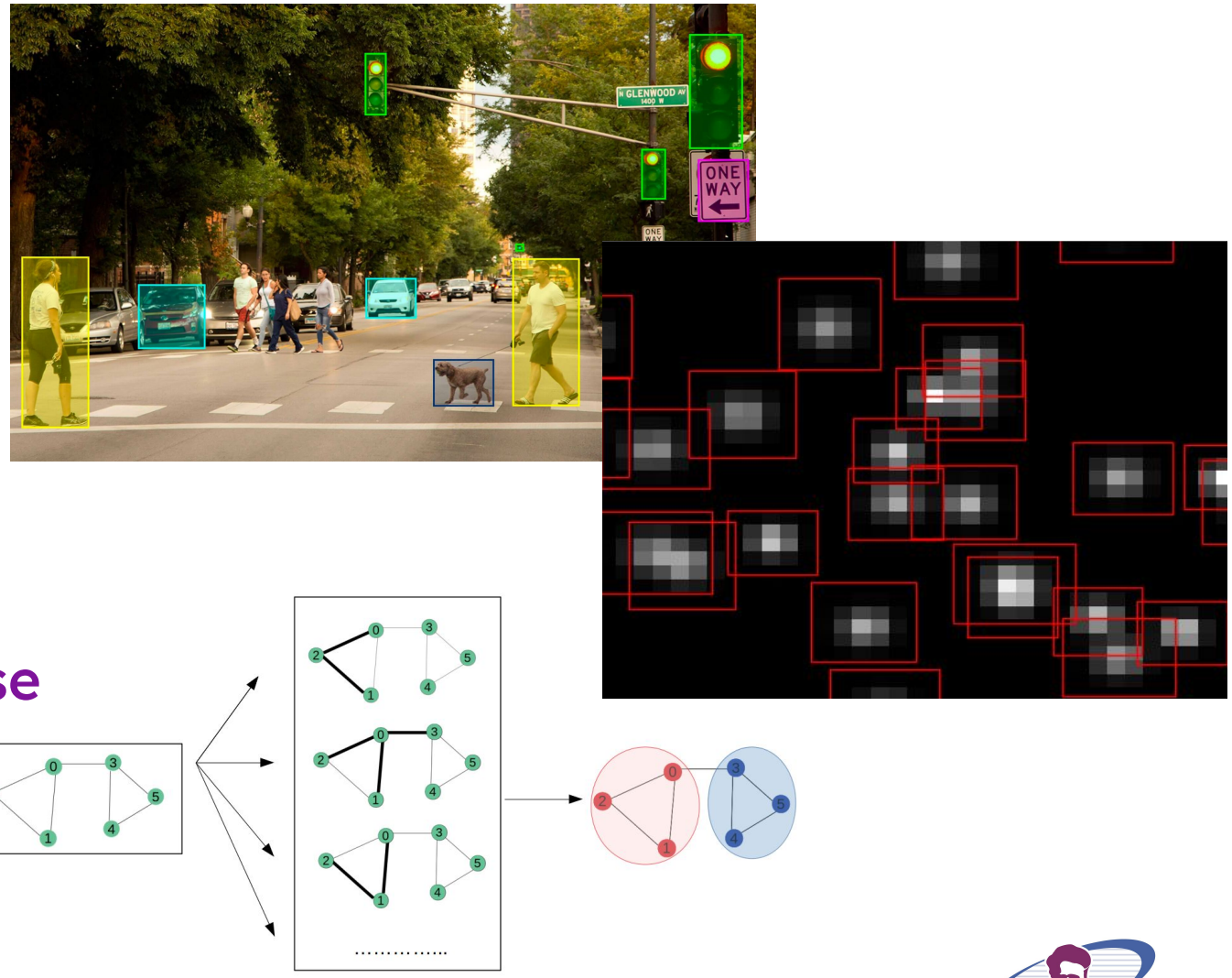
MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Initial Proposal

Two-step CNN + GNN

1. Transform detector response into eta-phi projection of calorimeter.
2. Use lightweight Region Proposal Network to identify Regions of Interest (ROIs).
3. Pass ROIs to GNN in original calorimeter coordinates.
4. Classify nodes in the graph as noise or members of a cluster.
5. (Incorporate splitter decisions, cluster classification and energy regression.)



SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Training and Secondment

Since beginning in October:

- Followed the Statistical Methods in Physics UniGe course.
- Attended ATLAS software week tutorial.
- Attending the CHIPP winter school January 2023.

Secondment at LightBox:

- October + January coordination meeting.
- Envisioned start data in April/May, both ML and statistical analyses of financial market data.
- Predictive models on time-series data.

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

7

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Career Aspirations

- Continue work at interface of physics and data science.
- Particular focus on ML methods - new architectures, new legacy applications.
- Reduce computational cost/inefficiency of physics reconstruction.
- Simplify decision making processes with explainable, easy-to-use algorithms.

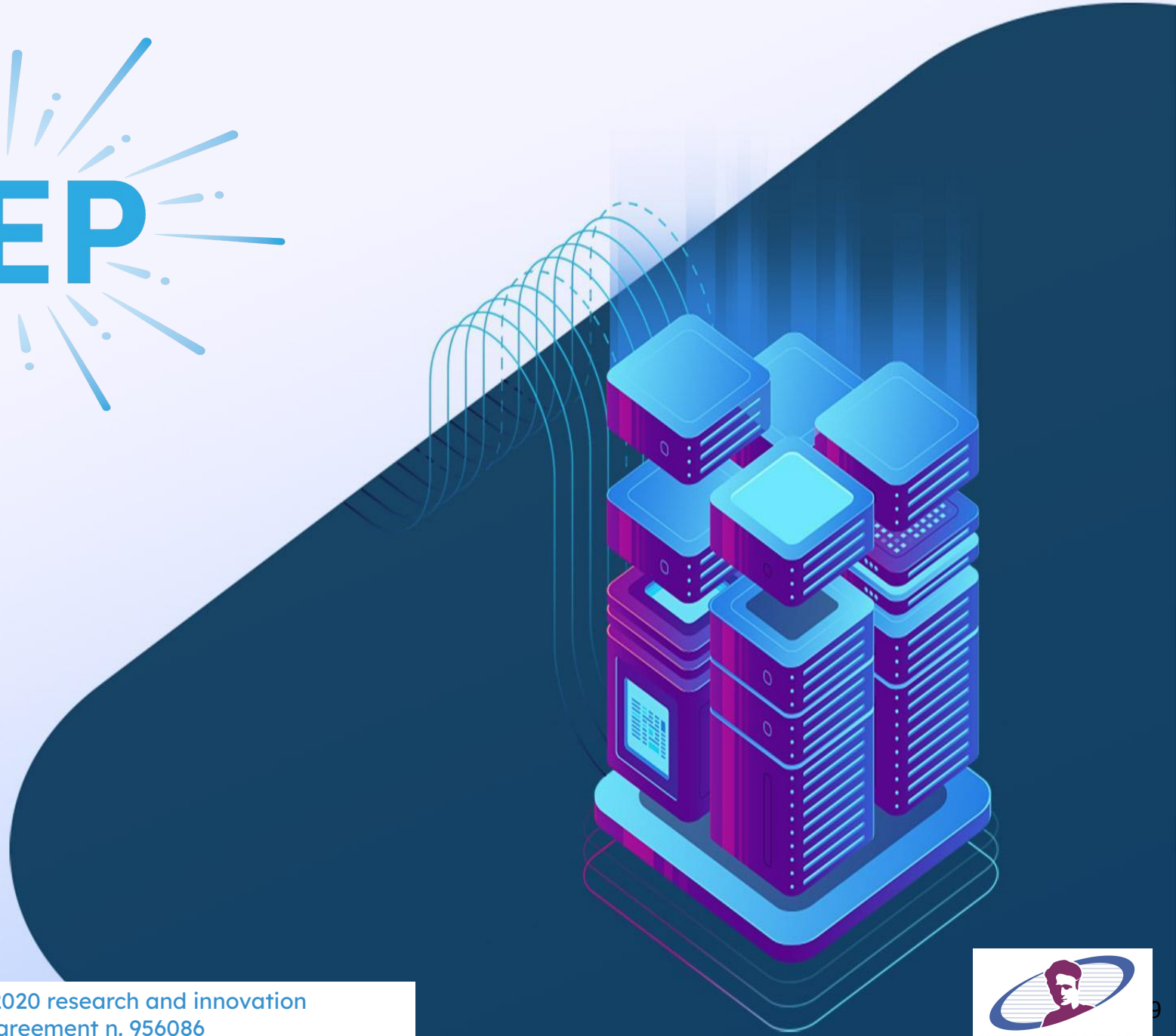
SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

8

SMARTHEP

REAL-TIME ANALYSIS FOR
SCIENCE AND INDUSTRY



SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

SMARTHEP is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2020, under Grant Agreement n. 956086

MARIE CURIE ACTIONS