Analysis Grand Challenge with ATLAS PHYSLITE data

IRIS-HEP FELLOW:

DENYS KLEKOTS

SUPERVISORS:

VANGELIS KOURLITIS

ALEXANDER HELD

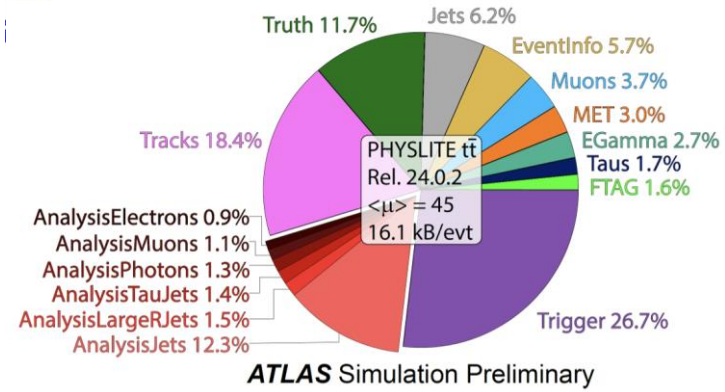
MATTHEW FEICKERT

Analysis Grand Challenge

- It is the benchmark to ensure that different pieces of software work fine together to give physically valuable results.
- Start from columnar data, including the processing of that data and construction of the statistical model with relevant visualisation at each step.
- The ATLAS AGC to be based on another already existing AGC with CMS data, find more here
- Usage of ATLAS Open Data collected in 2015 and stored in PHYSLITE format → testing the I/O of this data format.

PHYSLITE Data Format

1. Unskimmed and monolithic
2. Contains already calibrated objects
3. Objects are loosely preselected



Target size	MC	Data
PHYSLITE	12	10

kB per event

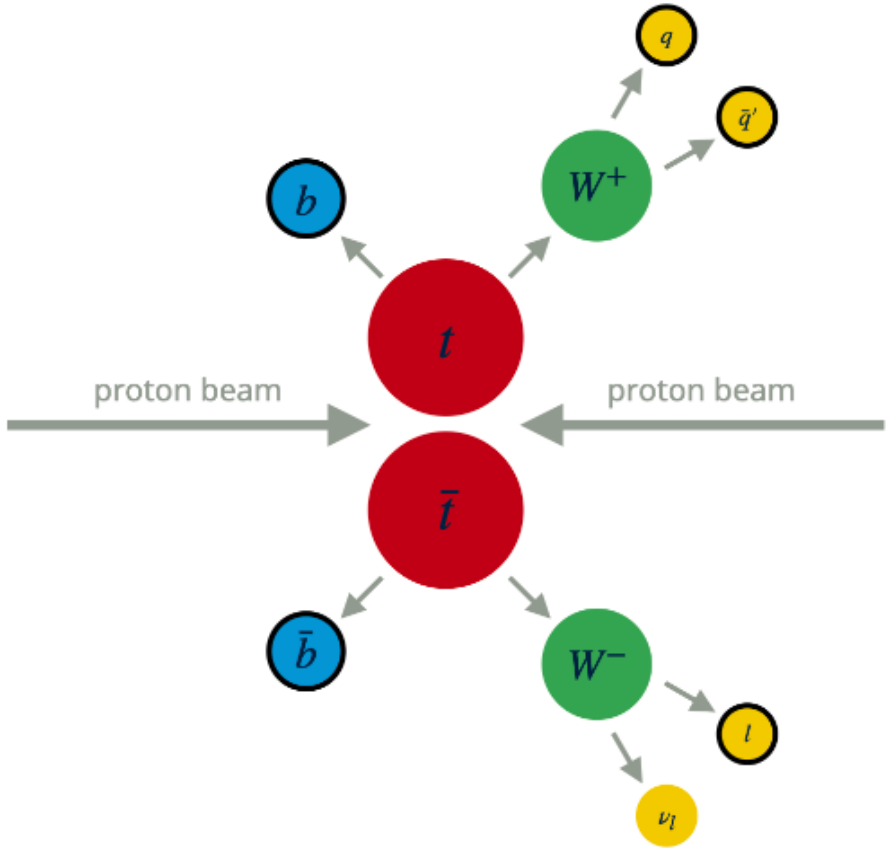
ATLAS releases the 2015+2016 physics proton–proton collision data in PHYSLITE format for research

Accompanied by “an appropriate set of simulated Monte Carlo samples”

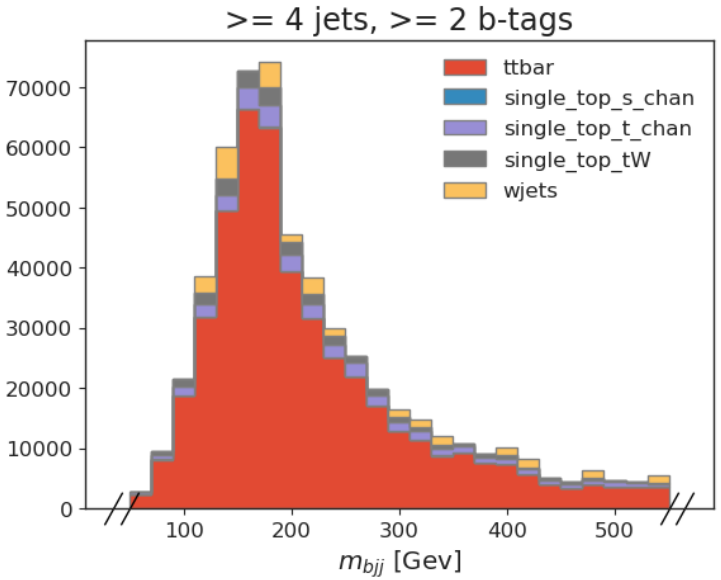
Distributed by opendata.cern.ch, support material at opendata.atlas.cern

$t\bar{t}$ pair data analysis specification

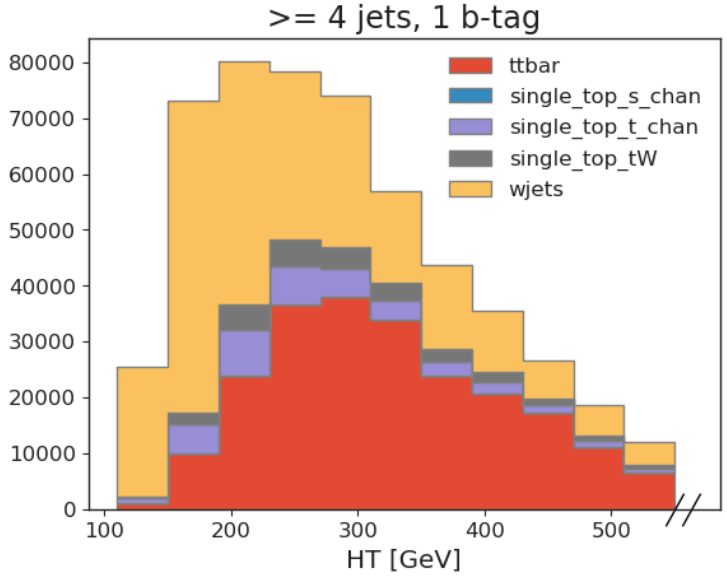
- The project implies top quark pair production data analysis.
- $t\bar{t}$ pairs are created in result of the proton-proton collision in the LHC.
- Event selection:
  - exactly one lepton
  - at least two b-tagged jets
  - selections on the transverse momentum of the jets and leptons



Observable physical results



Reconstructed mass of the t-quark



The scalar sum of the transverse momentum of the jets

Actual histograms shown here are from AGC performed by Alexander Held on CMS data

Timeline, key points and current progress

Timeline:

1. Become familiar with the software ecosystem;
2. Studying previous AGS with CMS data;
3. Develop code to work with PHYSLITE data format;
4. Implement the ATLAS PHYSLITE Analysis Grand Challenge;
5. Wrapup the project, prepare a markdown documentation;
6. Prepare slides and final presentation of the project.

Current situation:

- ATLAS Open Data in PHYSLITE format, measured in 2015, has not been published yet.
- Currently working with the MC-generated data to prepare code to be ready for the ATLAS data.

Project deliverable:

A public and reproducible AGC implementation.

Thank you for your attention