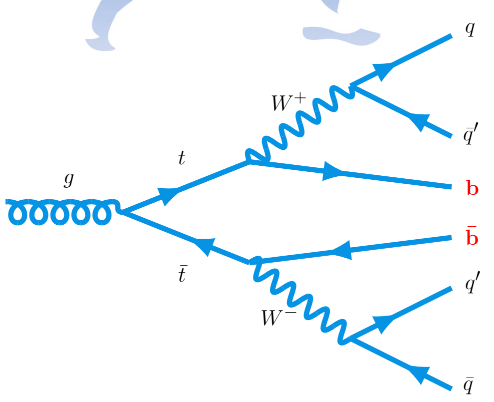
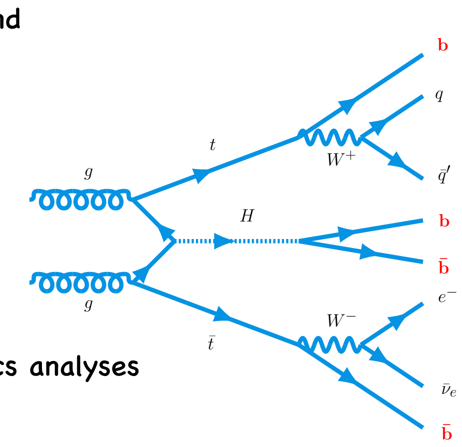
Optimisation of the ATLAS Deep Learning Flavour Tagging Algorithm

- LHCP 2020 Poster Session -

Why b-Tagging?

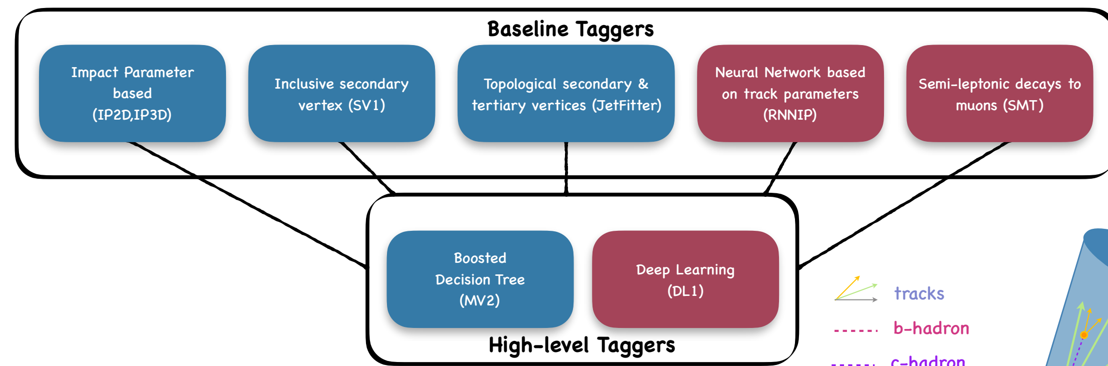
- Several interesting physics processes have b-quarks in their final state
- Or a veto on b-quarks can suppress the background

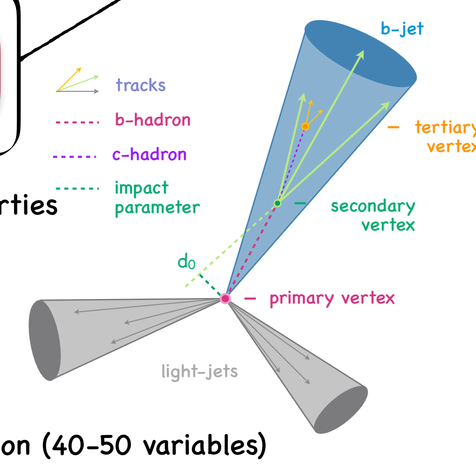




- Heavy-flavour tagging important tool for physics analyses
  - Precision measurements
  - Search for new physics

b-Tagging Structure in ATLAS



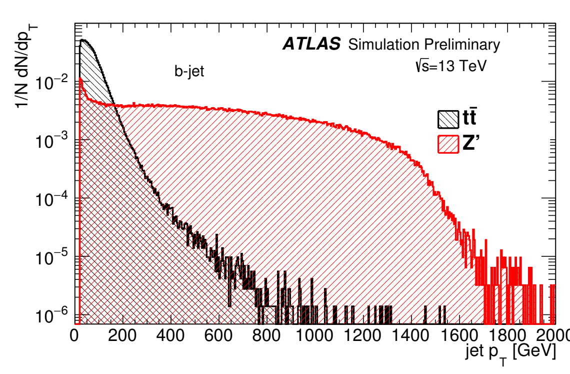


- Baseline taggers deploy specific heavy flavour jet properties
  - Long lifetime (~1.5 ps ➔ ~3mm track in detector)
  - High mass (~5 GeV)
  - High decay product multiplicity
  - b-hadron decays to a c-hadron ($|V_{cb}|>>|V_{ub}|$)
- High-level taggers (MV2 & DL1) combine these information (40-50 variables)

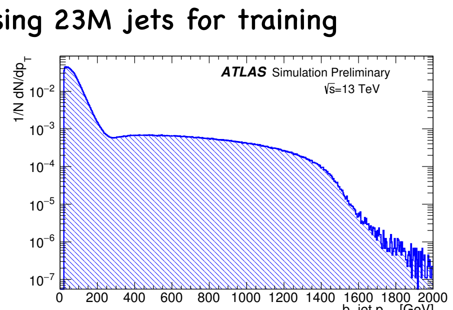
Training Samples

- Using hybrid sample composed of SM $t\bar{t}$ and $Z' \rightarrow q\bar{q}/b\bar{b}$ events

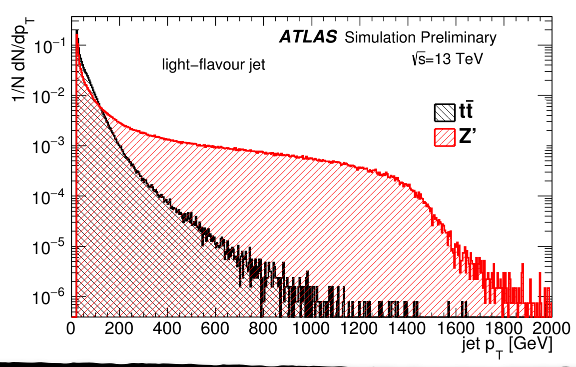
  ➔More statistics in higher p_T region
- Undersampling approach applied to match p_T and $|\eta|$ distributions for all 3 flavour categories
  - Ensure independency of tagging from kinematics
- Using 23M jets for training



ATL-PHYS-PUB-2017-013



ATL-PHYS-PUB-2017-013

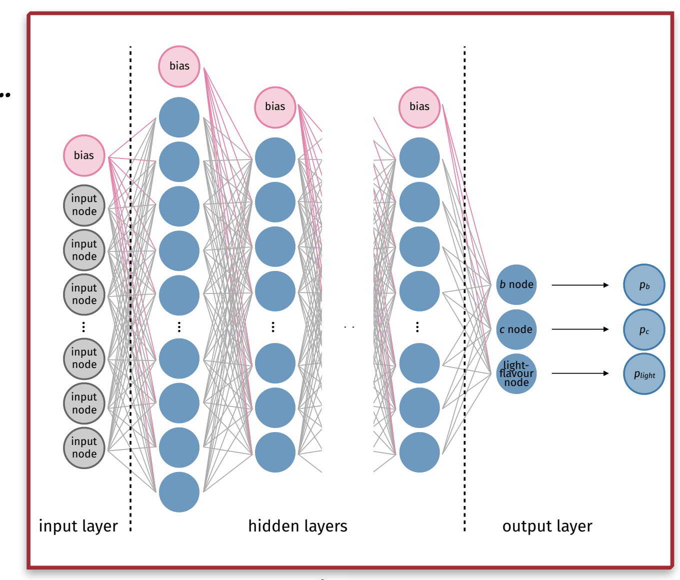


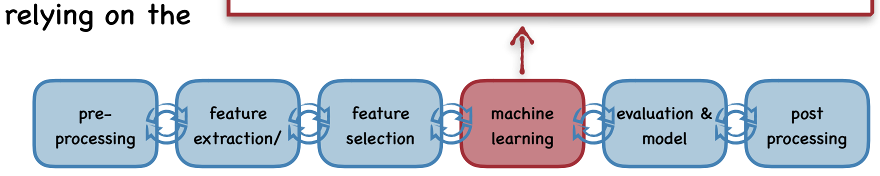
Deep Neural Network Architecture

- Deep neural network requires also
  - Preprocessing, feature selection, ...
- Network with fully connected layers
- Multi-class output ➔ allows also c-tagging

$$\mathrm{DL1}_{score} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1-f_c) \cdot p_{\text{light-flavour}}}\right)$$

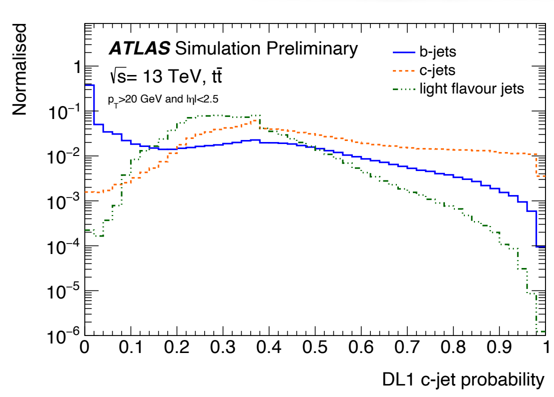
- Using Keras (2.2.4) framework with tensorflow backend
- Full training procedure relies on HDF5
- Application can be run in ATLAS reconstruction software, relying on the LWTNN C++ interface

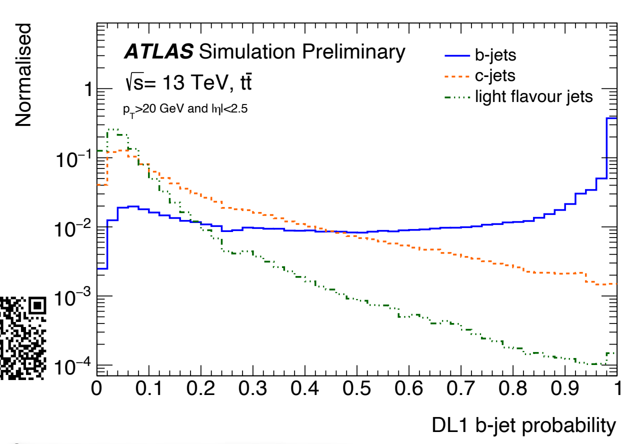




Network Output

- Each jet gets probability for being a b-, c- or light flavour jet

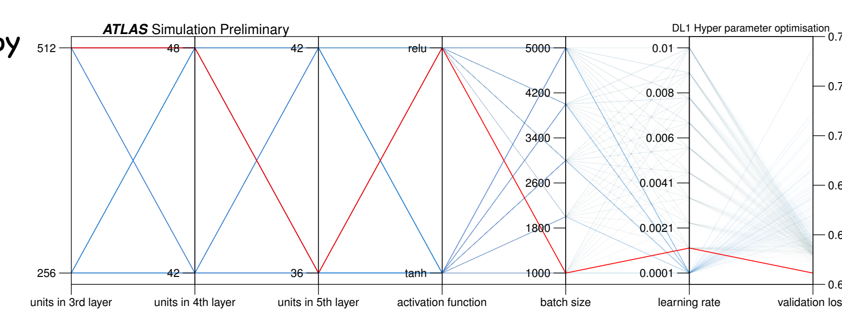


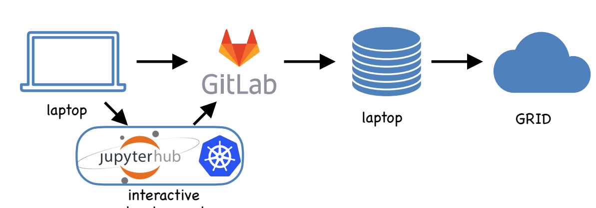


- Good separation of b- & light-jets
- b- & c-jets have more similar physics behaviour

Hyper Parameter (HP) Optimisation with GRID GPUs

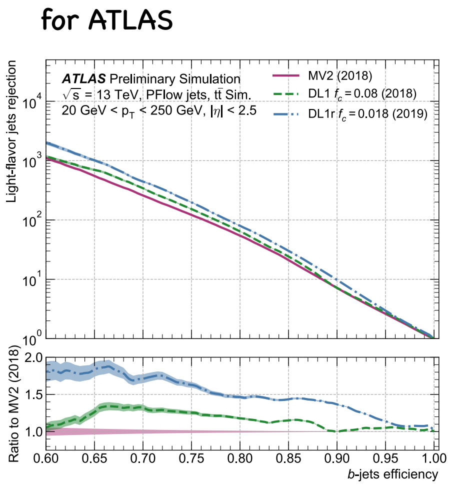
- Using docker image (built by Gitlab CI) for jobs
- Configurable amount of HP combinations (configs)
- Workflow optimised for GRID-submission
- 800 combinations over 5 HP dimensions
- Optimisation provides good results

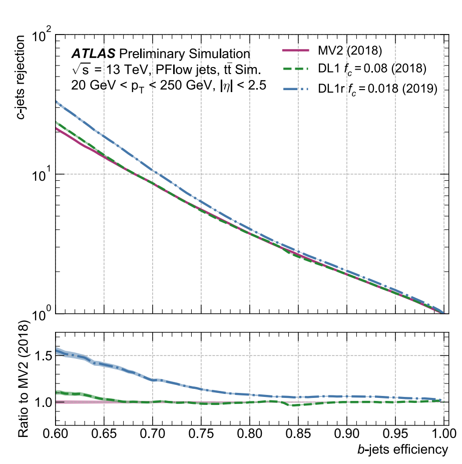




Final Training Results

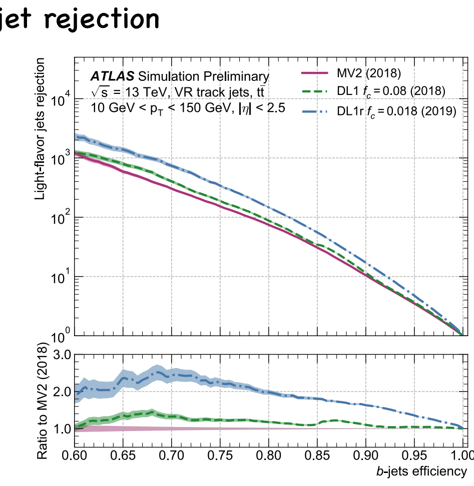
- Dedicated trainings of Particle-Flow jets and Variable Radius Track jets
  - New b-tagging recommendations for ATLAS

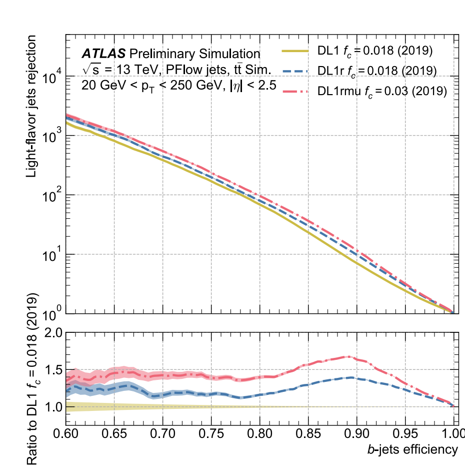




- Increase in performance up to 100% for light-jet rejection & up to 50% for c-jet rejection

- Performance gain similar for Particle-Flow and Variable Radius Track jets
- Gain in light-jet rejection higher than in c-jet rejection





- RNNIP (track information based NN) increases performance by ~25%
- Soft muon info shows another ~20% improvement (very difficult to calibrate)

Manuel Guth (Uni Freiburg, CEA Saclay), 28.05.2020
on behalf of the ATLAS Collaboration