

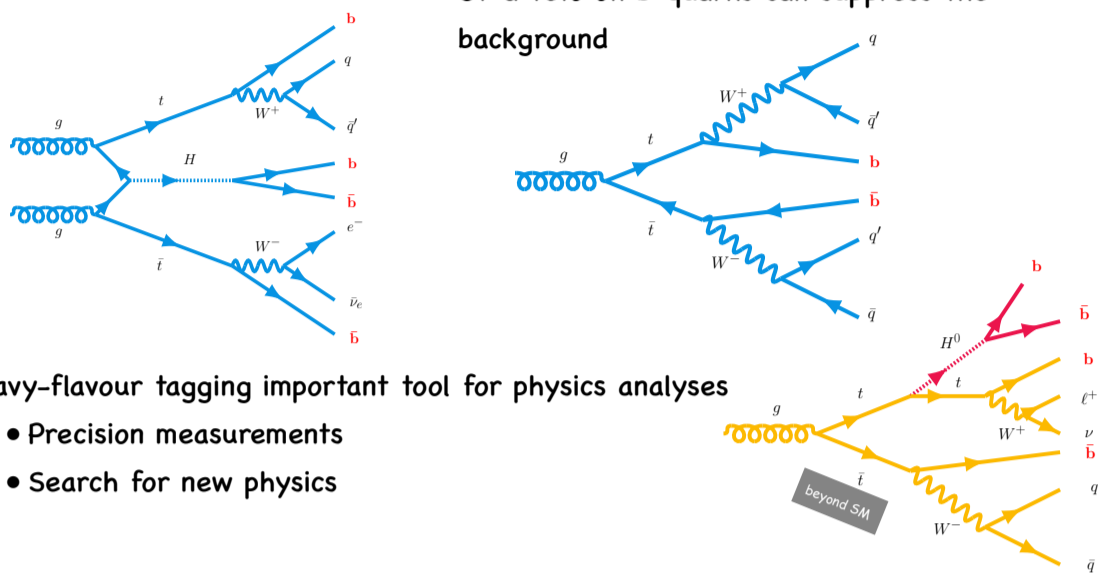
Advancement of the ATLAS Deep Learning Based High Level Flavour Tagger DL1

-LHCC Poster Session -



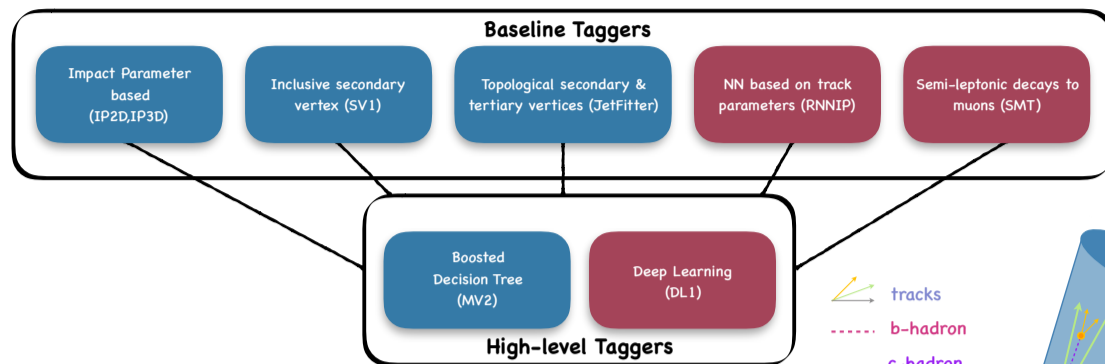
Why b-Tagging?

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

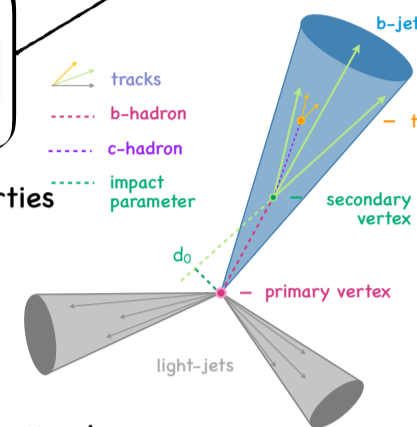


- Heavy-flavour tagging important tool for physics analyses
 - Precision measurements
 - Search for new physics

b-Tagging Structure in ATLAS

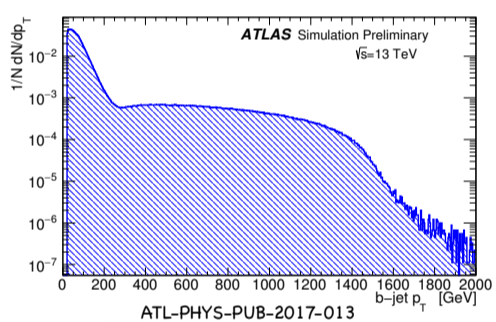
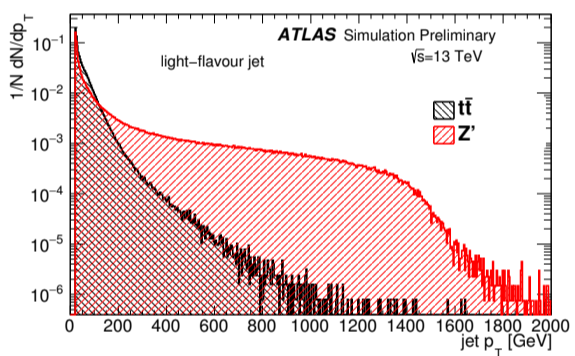
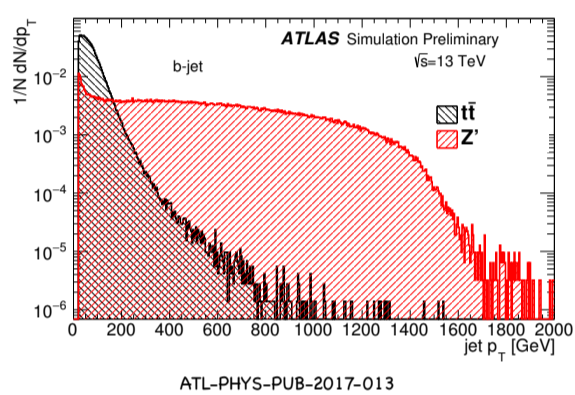


- Baseline taggers deploy specific heavy flavour jet properties
 - Long lifetime (~ 1.5 ps \rightarrow ~ 3 mm track in detector)
 - High mass (~ 5 GeV)
 - High decay product multiplicity
 - b-hadron decays to a c-hadron ($|V_{cb}| \gg |V_{ub}|$)
- High-level taggers (MV2 & DL1 [1]) combine these information (40-50 variables)



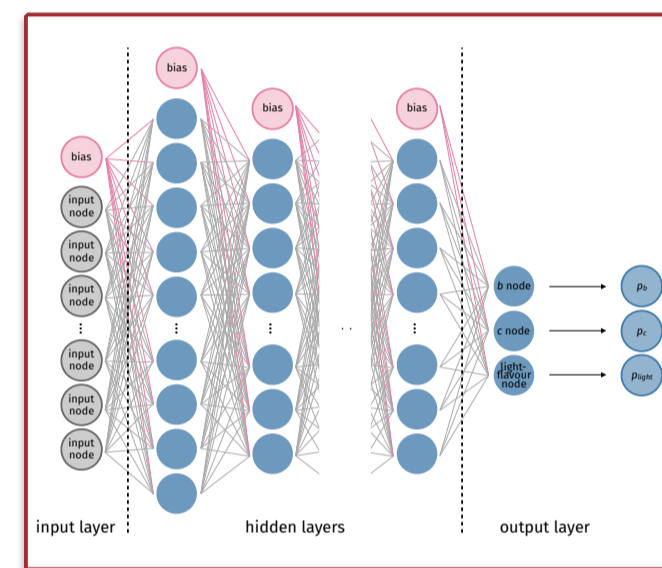
Training Samples

- Using hybrid sample composed of $t\bar{t}$ and Z' events
 - \rightarrow More statistics in higher p_T region
- Downsampling approach applied to match p_T and $|\eta|$ distributions for all 3 categories
 - Ensure independence of tagging from kinematics
- Different hybrid compositions tested



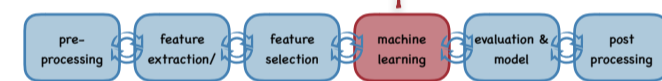
Deep Neural Network Architecture

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



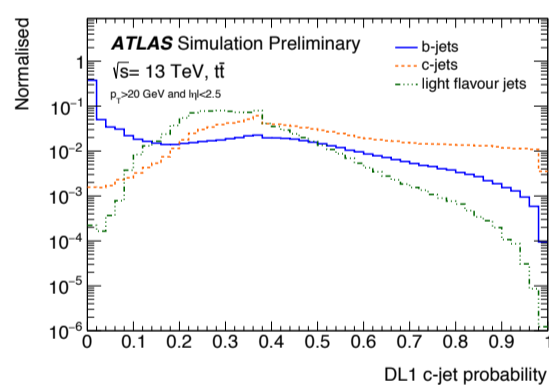
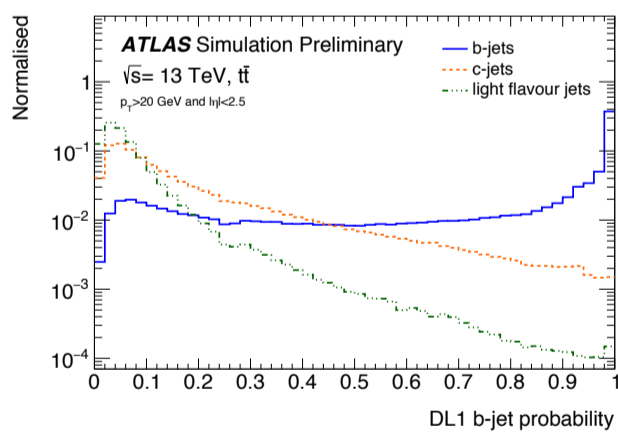
$$DL1_{score} = \ln \left(\frac{P_b}{f_c \cdot P_c + (1 - f_c) \cdot P_{light-flavour}} \right)$$

- 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 [2]



Network Output

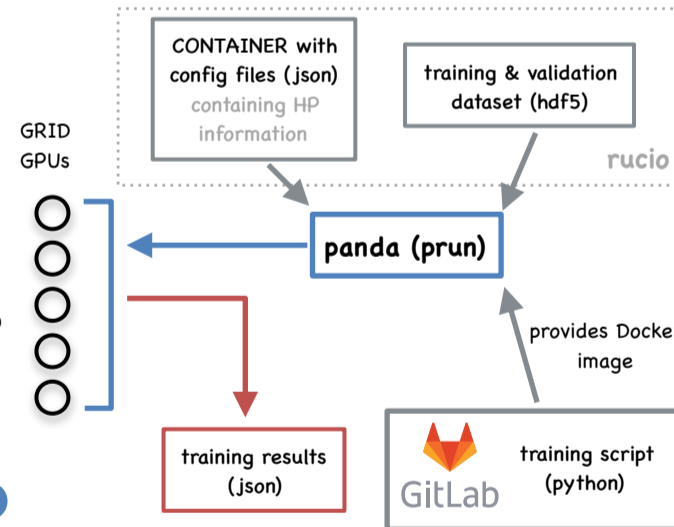
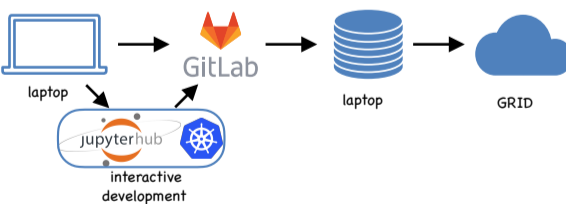
- Each jet gets probability for being a b-, c- or light flavour jet



Hyper Parameter (HP) Optimisation with GRID GPUs

- Using docker image (built by Gitlab CI) for jobs
- Configurable amount of HP combinations (configs)
- Using panda job splitting to distribute HP configurations over sites / jobs
- Interactive development on JupyterHub deployed with Kubernetes

Workflow



Optimisation Results

- 800 combinations over 5 HP dimensions tested
- Optimisation provides promising results
- Confirms relation between ROC curve and validation loss

