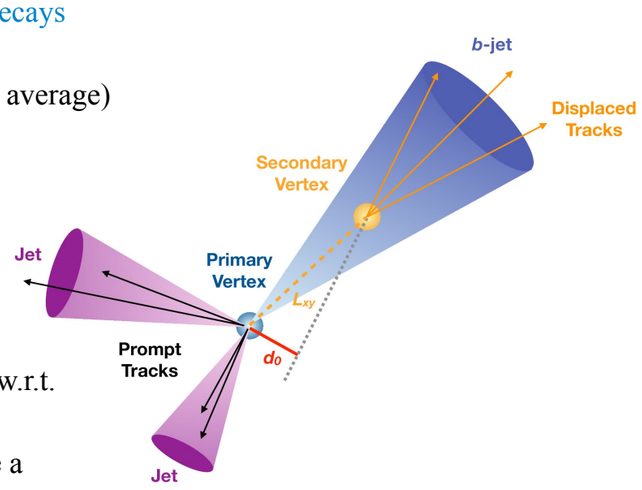
Flavour Tagging with Graph Neural Networks with the ATLAS Detector - EPS Poster Session -

Introduction:

- Identification of the jet flavour is a crucial task for many ATLAS analyses (e.g. $H \rightarrow b\bar{b}/c\bar{c}$).
- Improvements in the *b*-tagging performance bring better physics results.
- Graph neural networks and transformers open up a new era of performance for machine learning based algorithms

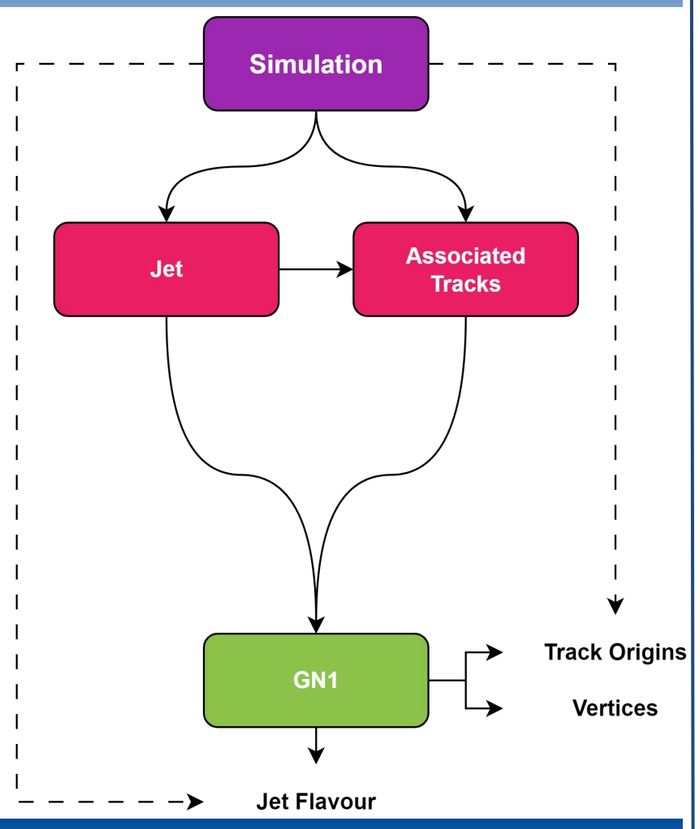
How does Flavour Tagging work?

- Using unique characteristics of *b*-hadron decays
  - Long lifetime (~1.5 ps → ~3 mm decay length in average)
  - High mass (~ 5 GeV)
  - Number of tracks per jet (~5 per jet for *b*-jets)
- These properties result in characteristic experimental signatures:
  - Displacement of secondary vertex w.r.t. primary vertex
  - Tracks from secondary vertex have a longer distance to primary vertex → Impact parameter d_0



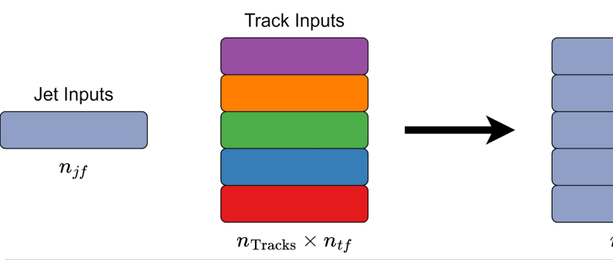
A New Hope - Graph Neural Networks

- ATLAS developed twp new graph neural networks (GN1/GN2) using one-stage approach
- Using concatenated information about jet and associated tracks directly
- No manual optimisation or training of a track-based neural network
- Less manual optimisation → Easier to maintain
- Training goal: Jet flavour classification
- Auxiliary tasks: Predict track origin and vertices



Inputs, Outputs and Preprocessing

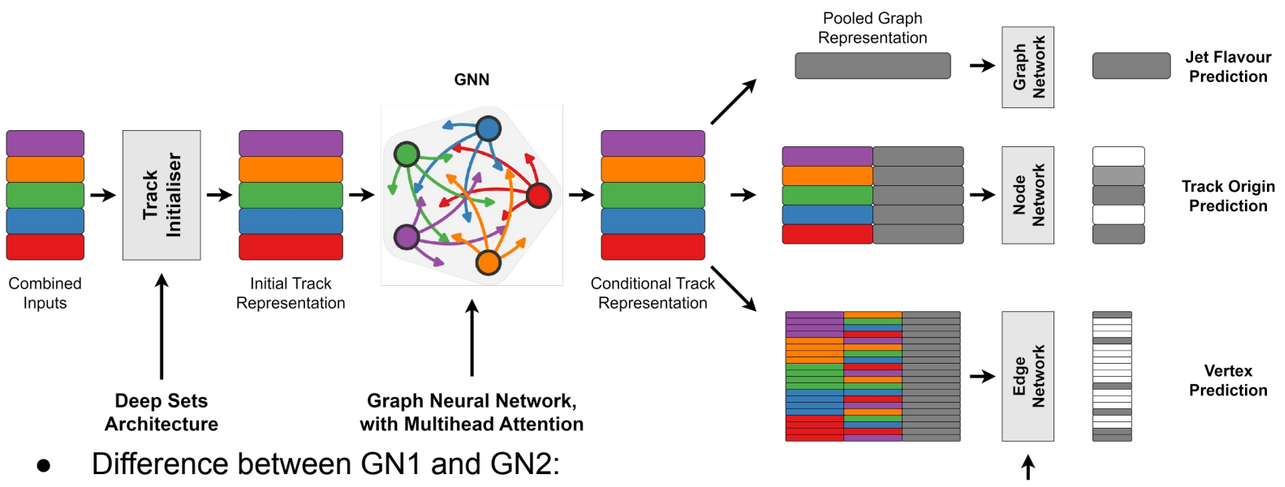
- Inputs:
  - Jet p_T and jet $|\eta|$
  - Track variables (e.g. impact parameters)
  - Hit information (e.g. number of pixel hits)
  - Using up to 40 associated tracks per jet
  - Concatenate jet information to each track
- Outputs:
  - p_b, p_c, p_u
  - Probabilities of jet originating from a *b*-, *c*- or light (*u, d, s*) quark
  - Using combination of the probabilities in a discriminant



$$D_b = \log\left(\frac{p_b}{f_c \cdot p_c + f_u \cdot p_u}\right)$$

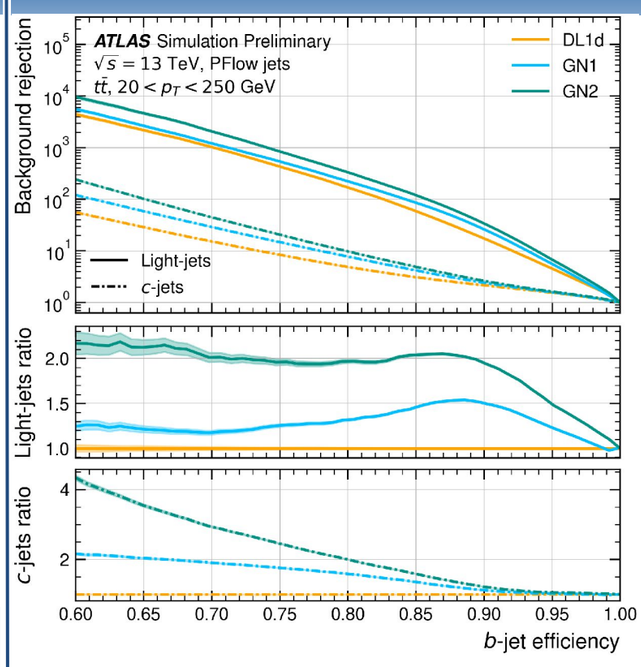
$$D_c = \log\left(\frac{p_c}{f_b \cdot p_b + f_u \cdot p_u}\right)$$

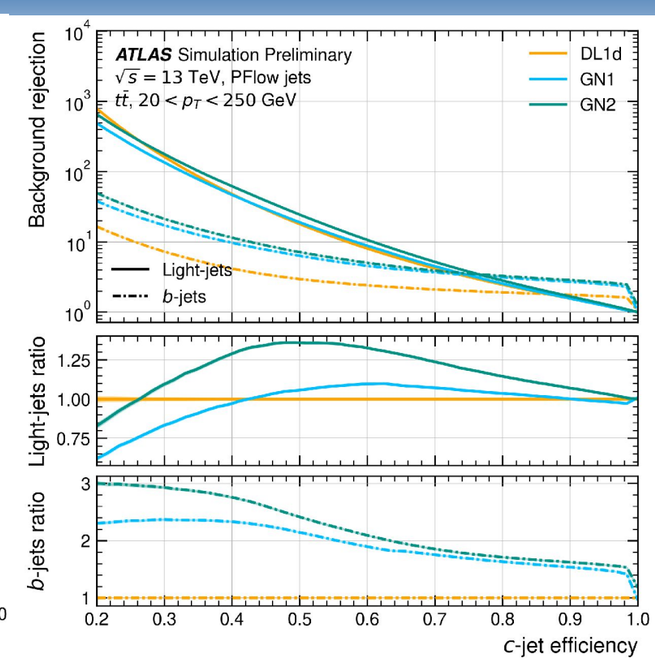
GN1/GN2 Architecture



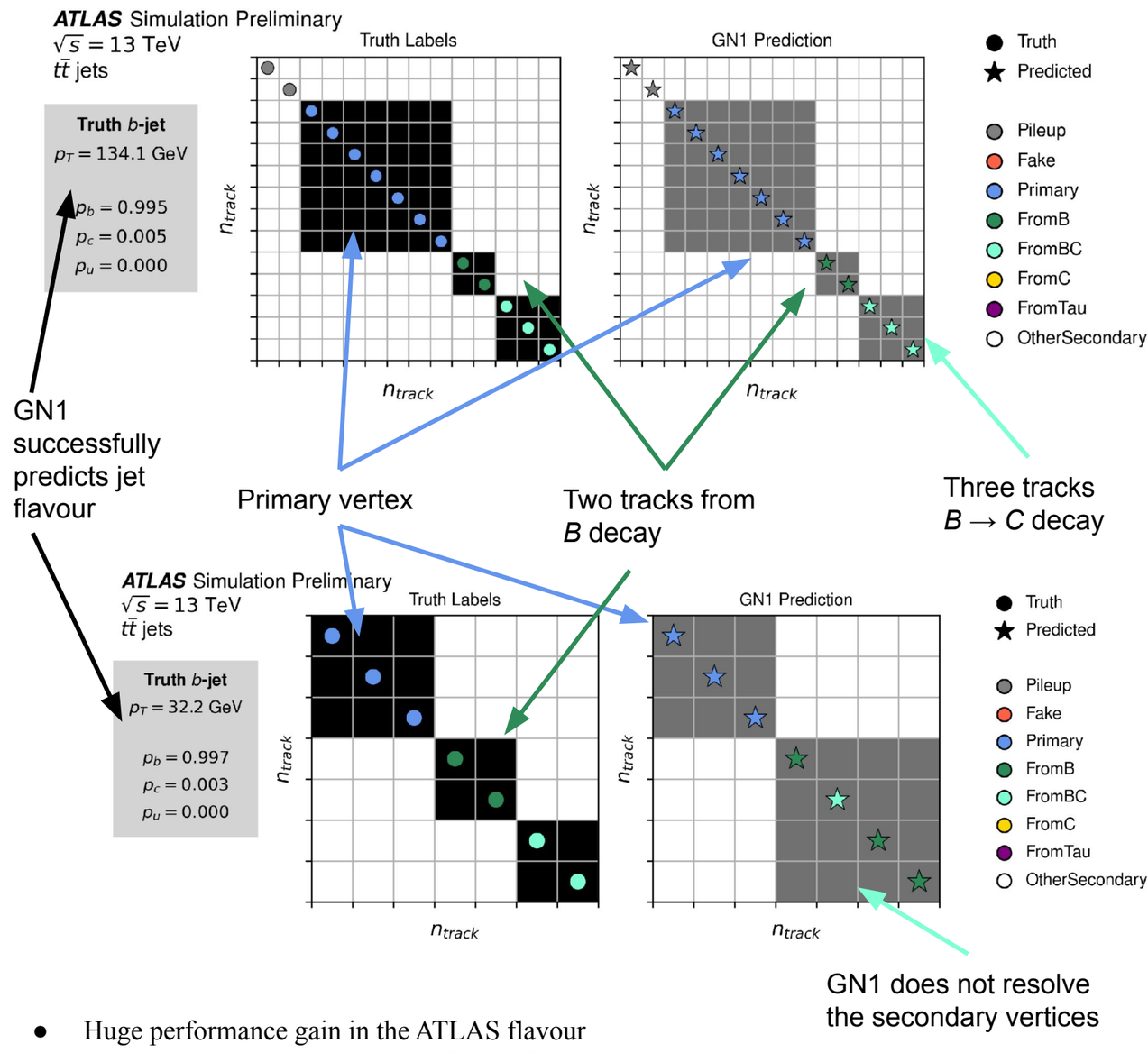
- Difference between GN1 and GN2:
  - Increased architecture size and training statistics
  - GN2 follows more closely the transformer architecture [arxiv:1706.03762]

Results of GN1/GN2

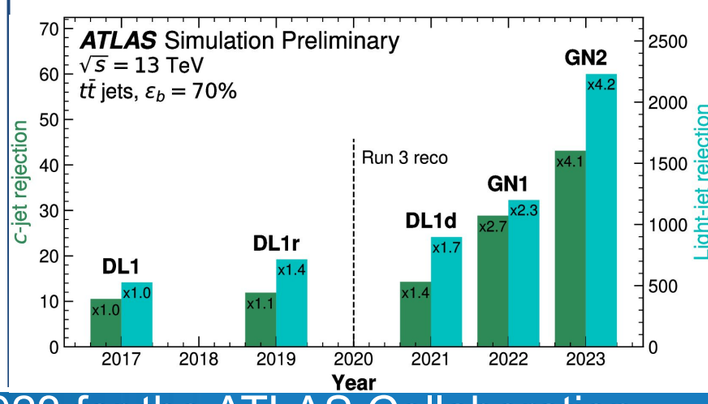


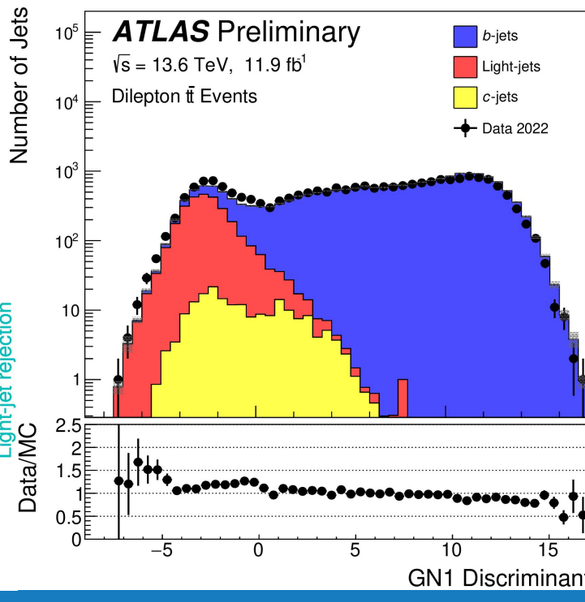


- Using the current default tagger for Run 3 (DL1d) as reference
- GN1/GN2 clearly outperform DL1d over the full range of *b*-efficiency for both *c*- and light jet rejection
- GN1/GN2 clearly outperform DL1d over the full range of *c*-efficiency in *b*- jet rejection. Also improvement in light-jet rejection!



- Huge performance gain in the ATLAS flavour tagging program due to new graph neural networks
- Data/MC agreement looks also very promising!
- For more info, check out the EP-IT Data Science Seminar talk about GN1/GN2





Alexander Froch (Albert-Ludwigs Universität Freiburg), 22.08.2023 for the ATLAS Collaboration

Plots taken from: Jet Flavour Tagging With GN1 and DL1d. Generator dependence, Run 2 and Run 3 data agreement studies, FTAG-2023-01
Configuration and performance of the ATLAS *b*-jet triggers in Run 2, Eur. Phys. J. C 81 (2021) 1087