HF jets analysis

16.03.2020 ALICE@IFJ meeting

Sebastian Bysiak

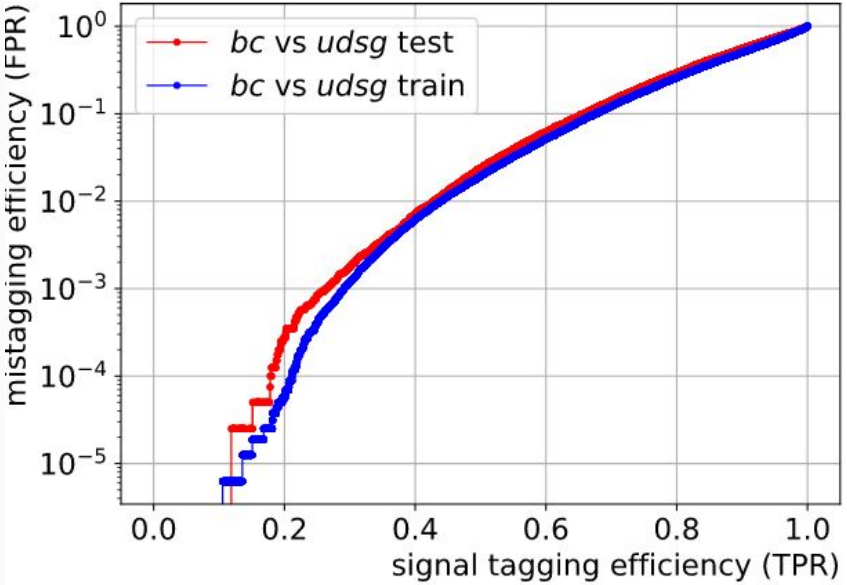
Outline

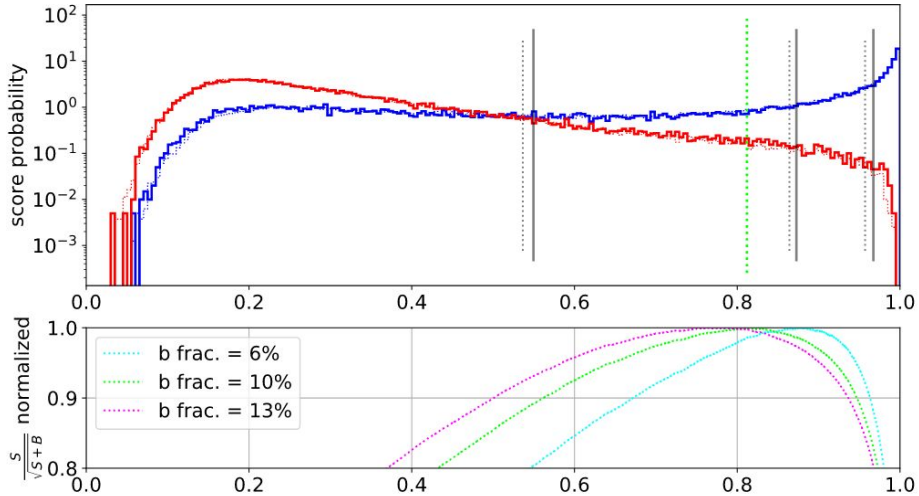
1. Progress in analysis
    - results after adding c-jets
    - performance comparison with others
2. Questions & issues
3. Plans for next week

Outline

1. Progress in analysis
    - **results after adding c-jets**
    - performance comparison with others
2. Questions & issues
3. Plans for next week

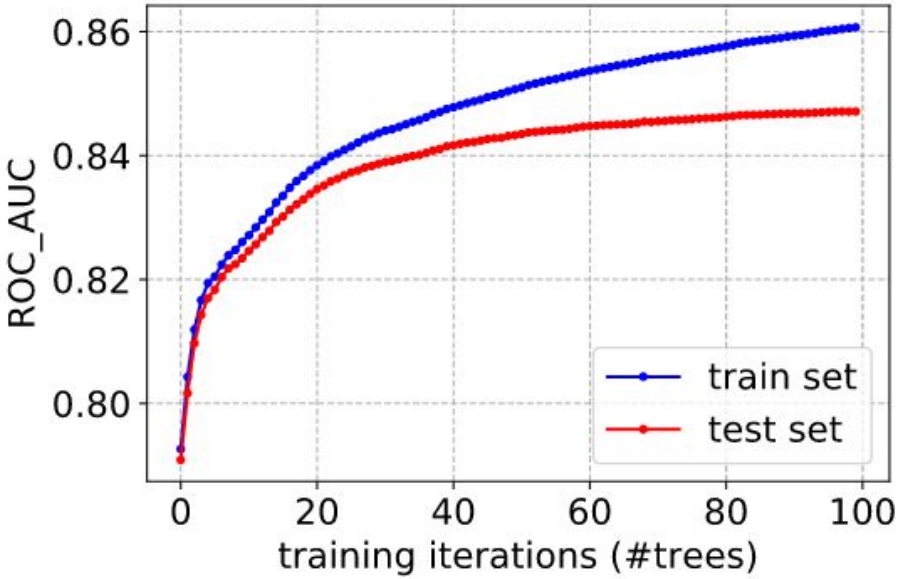
Results with c-jets: **bc-vs-udsg**





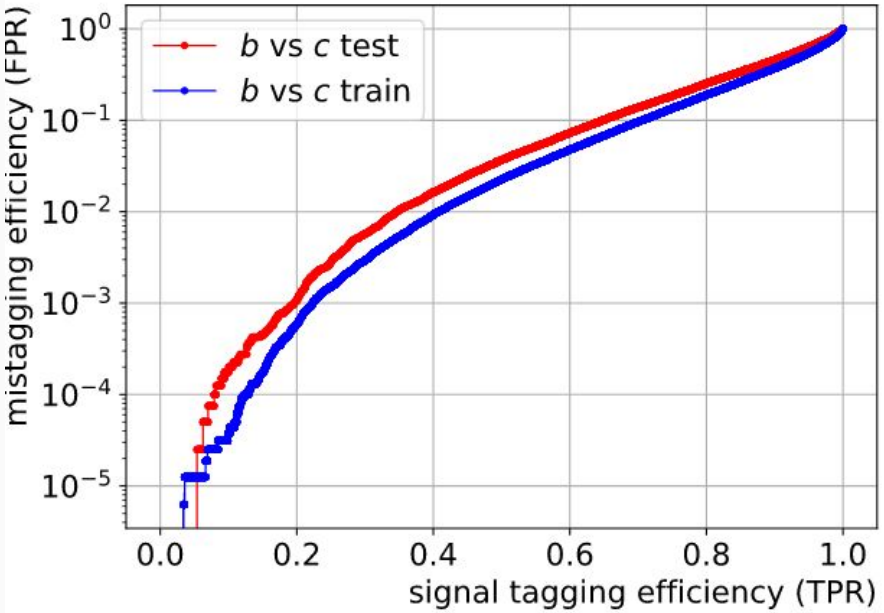
https://www.comet.ml/phd/bc-vs-udsg/bcf99db8f5a94b2184e6e13161c50bbe?experiment-tab=images

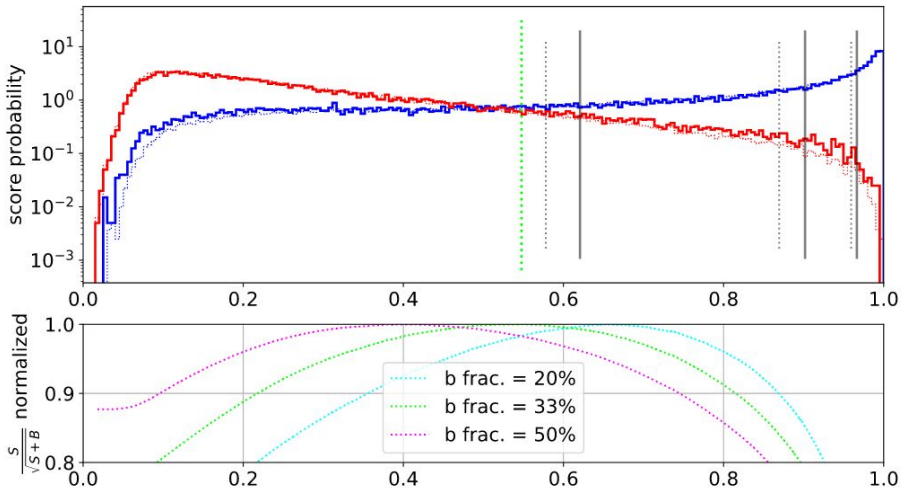
Results with c-jets: **bc-vs-udsg**



https://www.comet.ml/phd/bc-vs-udsg/bcf99db8f5a94b2184e6e13161c50bbe?experiment-tab=images

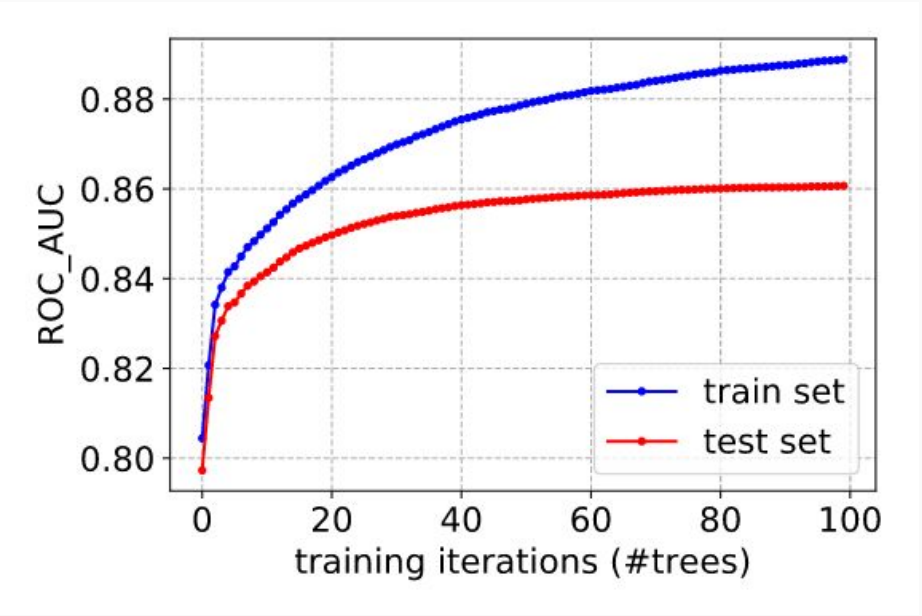
Results with c-jets: **b-vs-c**

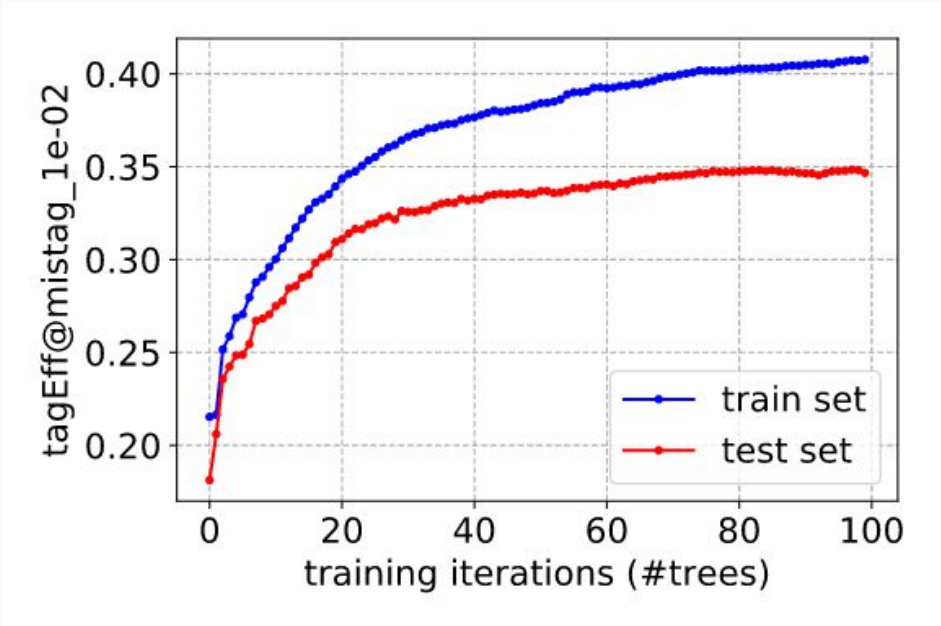




https://www.comet.ml/phd/b-vs-c/5c6cc4dfcd2743b99b39b2140075bde7?experiment-tab=images

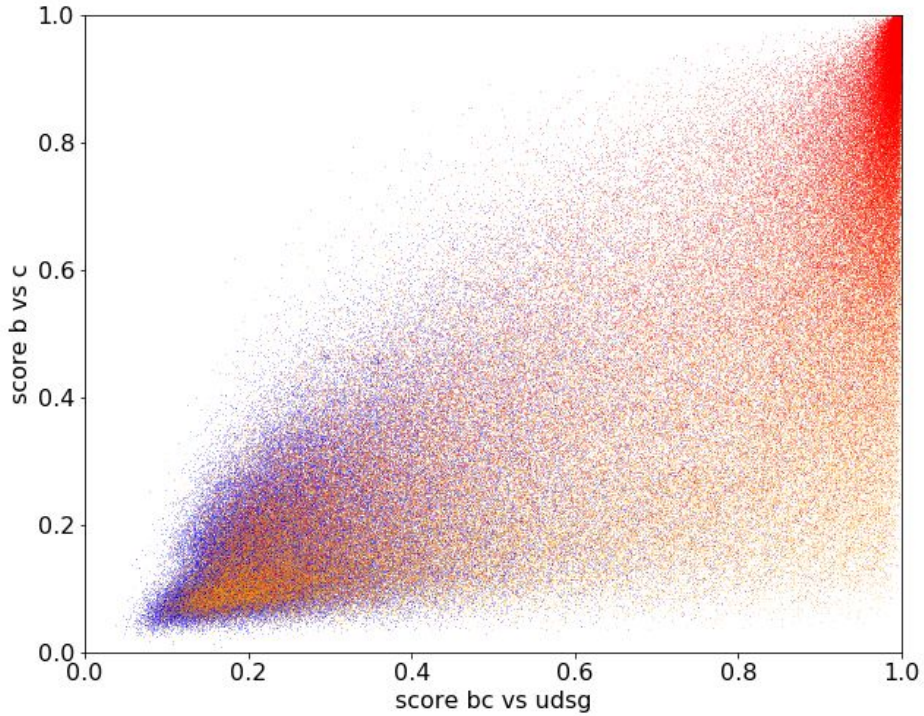
Results with c-jets: **b-vs-c**





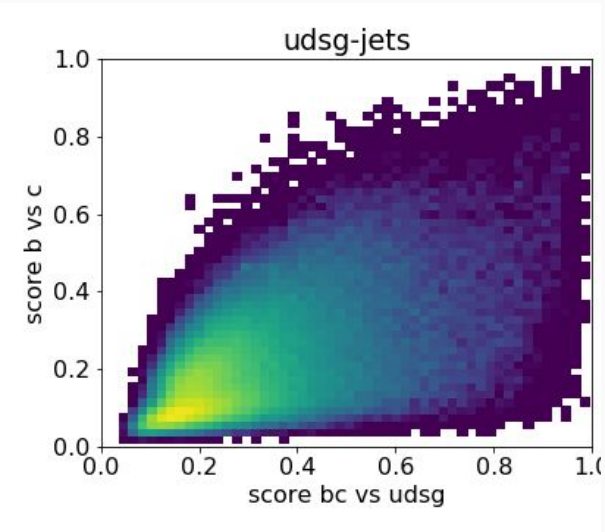
https://www.comet.ml/phd/b-vs-c/5c6cc4dfcd2743b99b39b2140075bde7?experiment-tab=images

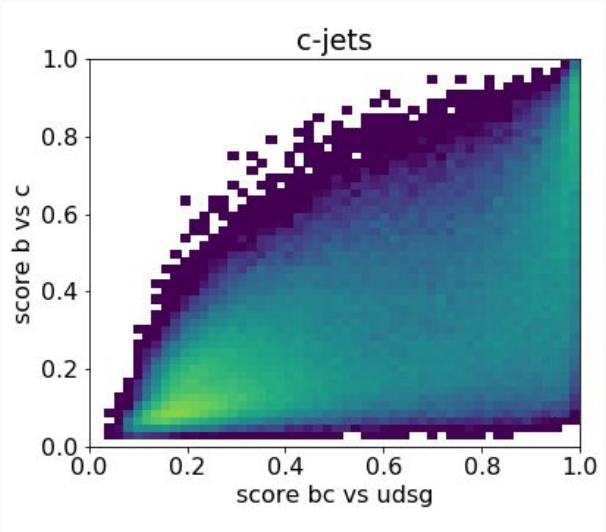
Results with c-jets: **b-vs-c-vs-udsg**

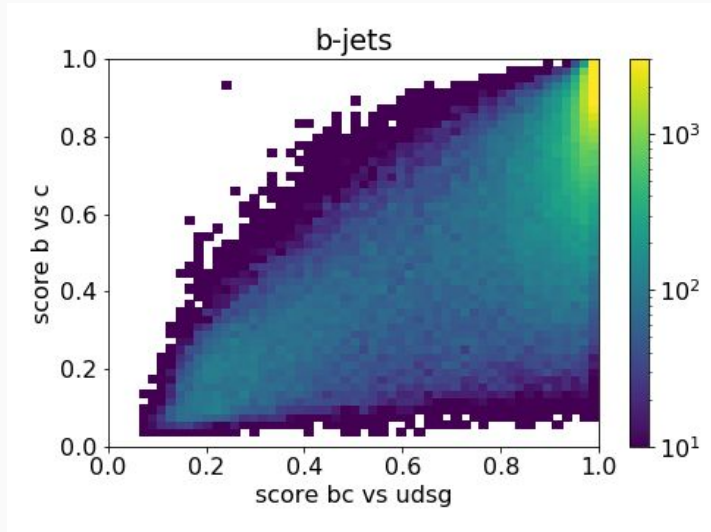


b-jets
c-jets
udsg-jets

Results with c-jets: **b-vs-c-vs-udsg**



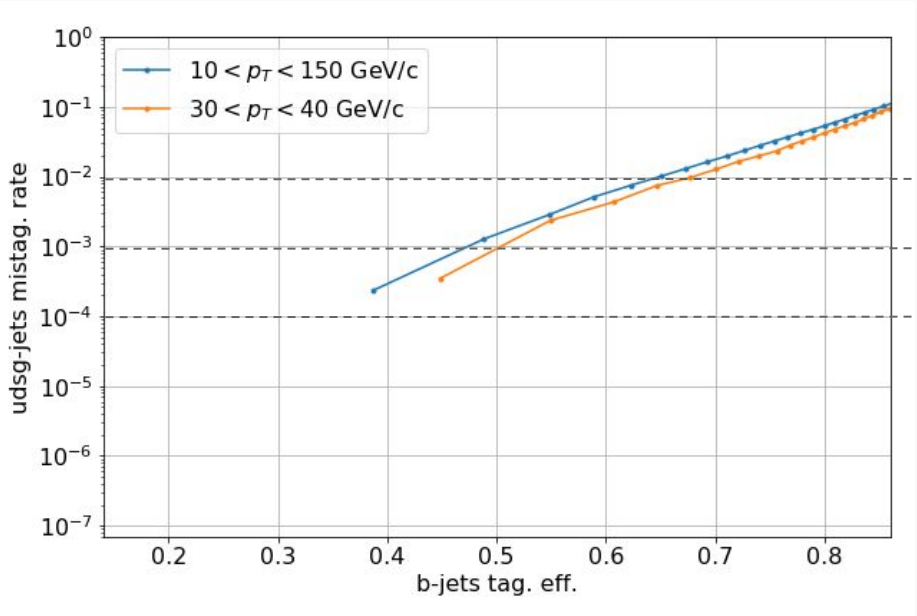


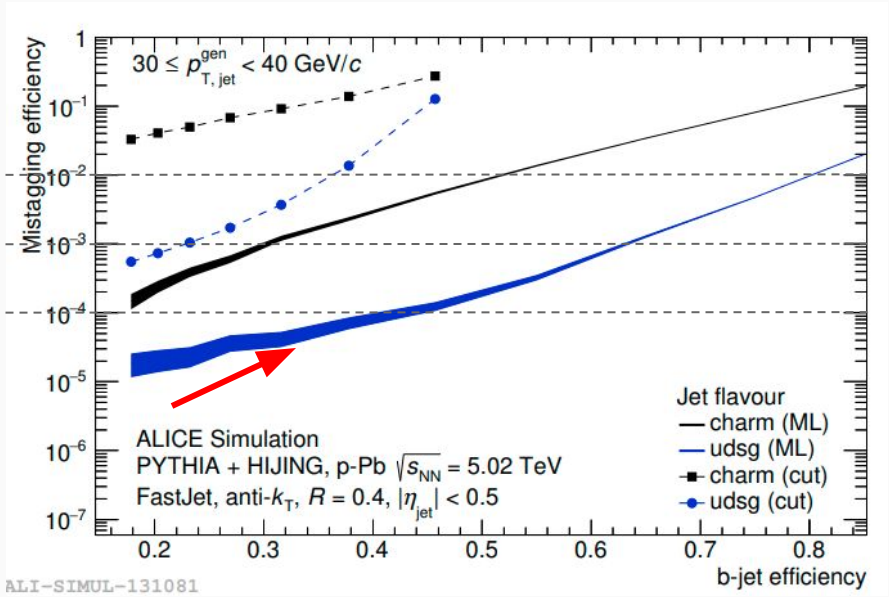


Outline

1. Progress in analysis
    - results after adding c-jets
    - **performance comparison with others**
2. Questions & issues
3. Plans for next week

Performance comparison - Rudiger

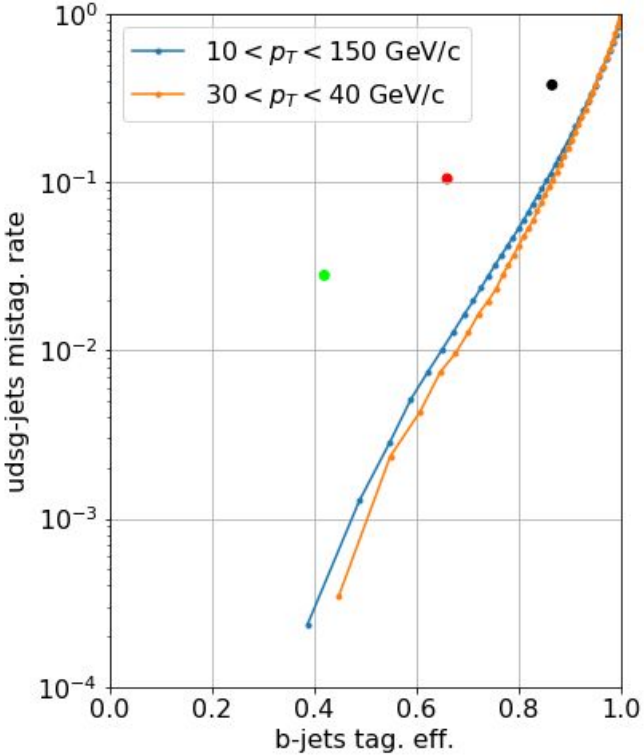




- pp vs pPb
- c-jets treated as same class as: b-jets vs udsg-jets

Fig. 2 in
https://arxiv.org/pdf/1709.08497.pdf

Performance comparison - Hadi Hassan



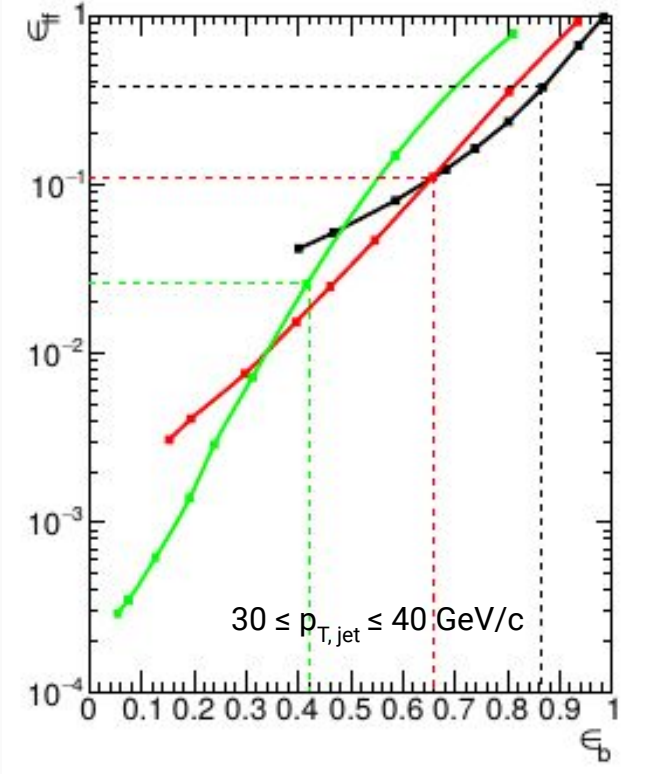
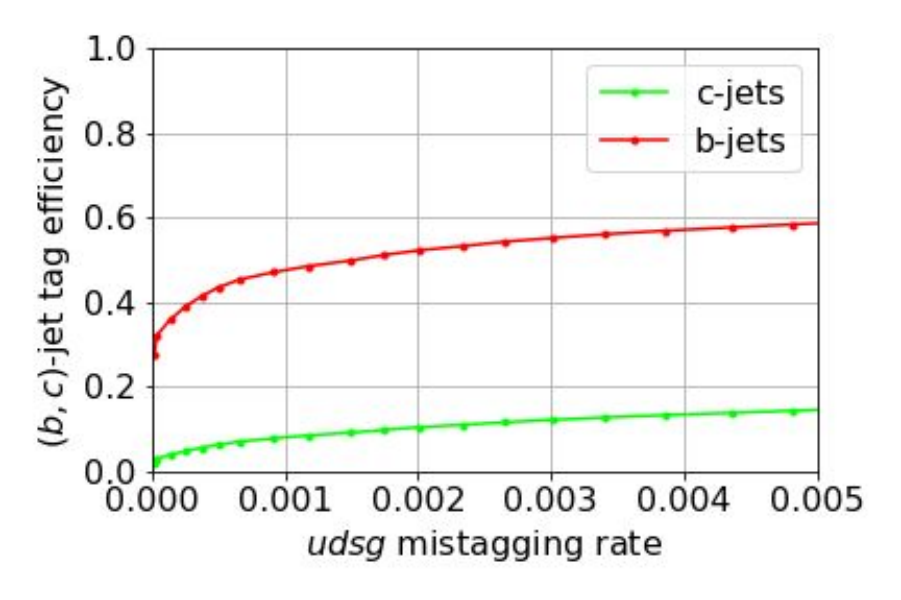


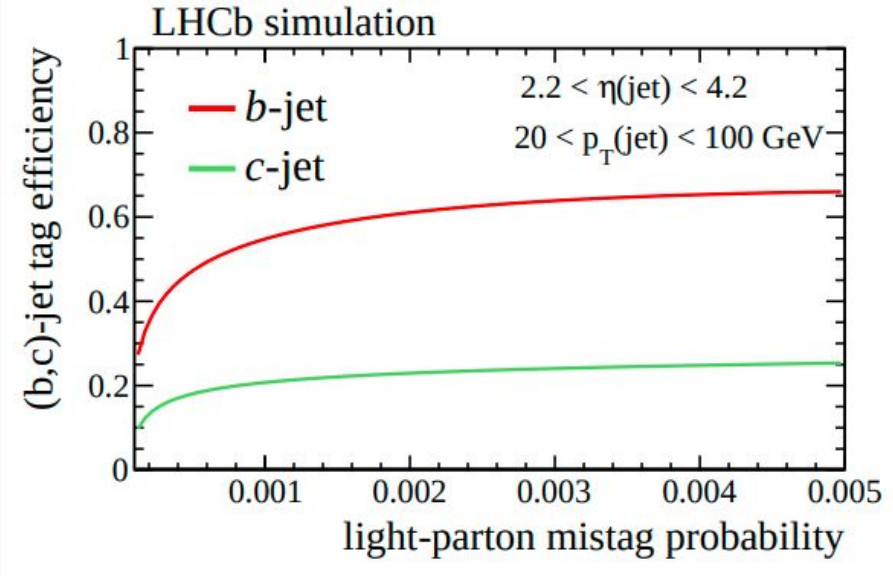
Fig. 10 in
https://alice-notes.web.cern.ch/system/files/notes/analysis/982/2019-10-06-ALICE_analysis_note.pdf

- different MC productions (here LHC17pq)
- different selection criteria
- complex vs simple algo

Performance comparison - LHCb

Fig. 3 in https://arxiv.org/pdf/1504.07670.pdf





- different experiment

Plans for the next week

1. Apply on data:
    - 1D & 2D score distribution
    - check b- and c-jets fractions (vs pT)
    - train classifier: MC vs data
2. Design selection criteria for SV (just sort by chi2 ?), at least vizualize sth

- check available MC prod. for LHC17p, LHC17q -- hard pt-bins, hf-enhanced
- model improving (ML-side): PCA before BDT, feat. eng., incl. jet shapes, vary sorting, N_tracks and N_Sv
- model improving (physics-side):
  technically easy: PID (e.g. e- and its energy)
  technically hard: Lund diagram, D and B meson reconstruction

BACKUP

Metrics

I will report ROC AUC,
0.01 difference in ROC AUC corresponds to

- 5% eff. @ mistagging rate = 0.001
- 5.6% eff. @ mistagging rate = 0.01
- 2.6% eff. @ mistagging rate = 0.1

to be confirmed how general are these dependencies

c/incl. ratio = 2(4)-10% by https://arxiv.org/pdf/1905.02510.pdf

c/incl.jets ratio = 5-9% by Hadi Hassan

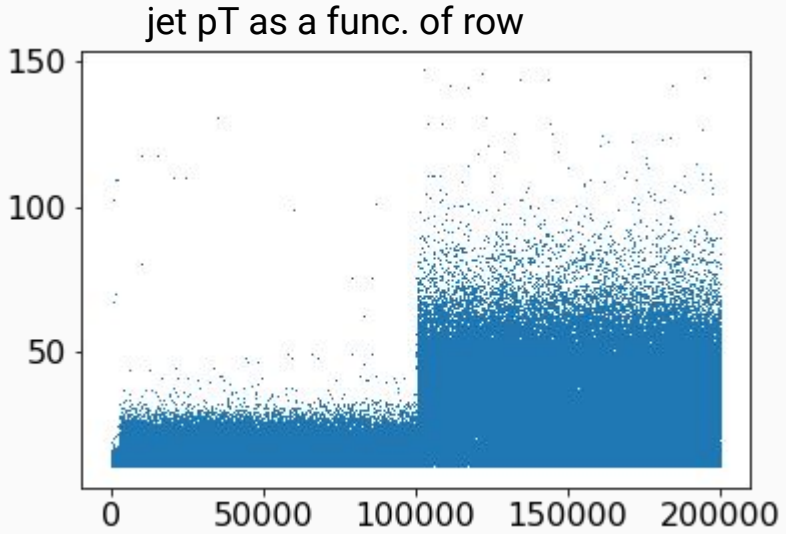
b/inc. jets ratio = 1-4% by Hadi Hassan

Problem with unshuffled hard-pt-bins

I was reading first N rows from my csv files

It appeared that they were filled in certain order: hard-ptbin after hard-ptbin, that's why jets with different pT were not distributed equally in rows, but rather first rows were mostly populated with low-pT jets

It affects experiments performed until 13.03.2020



lower_edges=(5 7 9 12 16 21 28 36 45 57 | 70 85 99 115 132 150 169 190 212 235)
higher_edges=(7 9 12 16 21 28 36 45 57 70 | 85 99 115 132 150 169 190 212 235 -1)

momentum dispersion: $p_T D = \frac{\sqrt{\sum_{i \in jet} p_{T,i}^2}}{\sum_{i \in jet} p_{T,i}}$.

angularity: