

ANOMALY DETECTION FULLY HADRONIC STATUS 5' REPORT

ANTONIO D'AVANZO ON BEHALF OF THE ANALYSIS TEAM

[Analysis twiki](#)

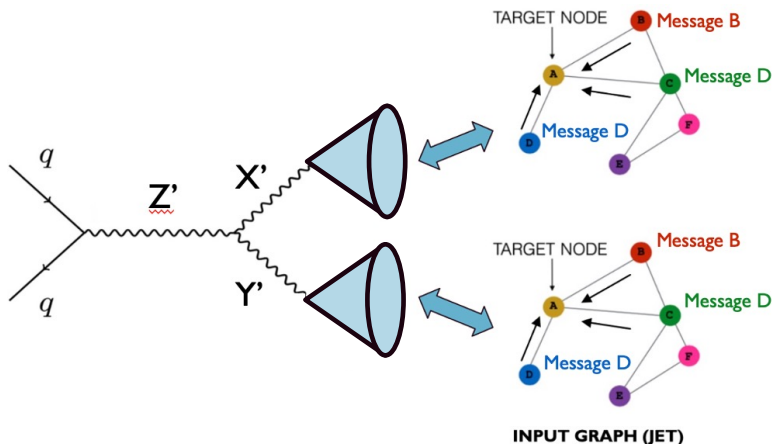
[Analysis glance](#)

[atlas-phys-hdbs-AD-JJ mailing list](#)

DBL Subgroup Meeting, CERN, 10/04/2024

THE ANALYSIS


- **Anomaly Detection** in fully hadronic events with message passing based Graph Neural Networks (GNNs).
- Graphs representing the final states jets, with 2 pT leading jets per event, built from transformed constituents.



- **Final goal:** Run 3 fully hadronic search
 - Completely model agnostic, 2 large-R jets per event
 - Signal region based on Anomaly Score cut.

✓ **toy model**

- ▶ R&D LHC Olympics dataset
- ▶ QCD dijet events as background
- ▶ $W' \rightarrow XY \rightarrow qq\bar{q}\bar{q}$ signal events
- ▶ $m_{W'} = 3.5\text{TeV}, m_X = 500\text{GeV}, m_Y = 100\text{GeV}$
- ▶ reconstructed with anti- k_T with $R = 1.0$

 **benchmarks**

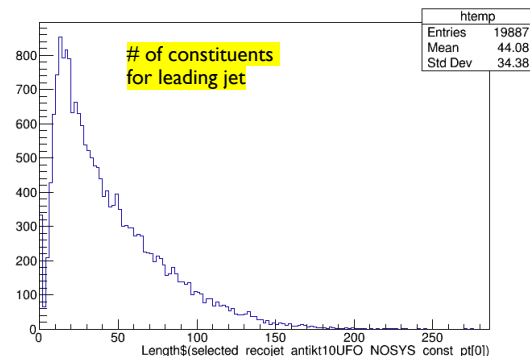
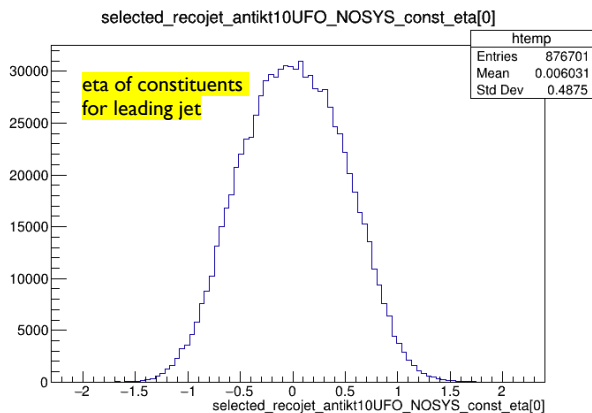
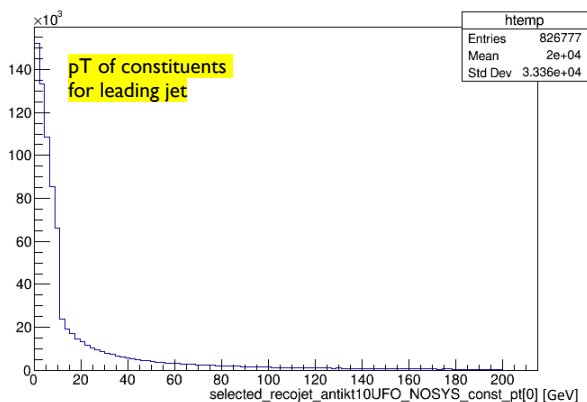
ML tasks

- ▶ **supervised approach:** RF, transformers, GNN for graph-classification
- ▶ **unsupervised approach:** GNNs, transformer + AE

More details in Graziella's last [status update](#)

BIGGEST WIN: NTUPLE MAKER!

- Production of ntuples from **our** run 3 [LLJ1 DxAOD](#) (DxAOD PHYS + constituents) based on [EasyJet](#) framework.
- Achievements:
 - Disabling b-tagging on large-R jets;
 - Customization of list of applied triggers;
 - Possibility to include additional variables computed from those already in DxAOD, in particular for jets.
 - **Addition of constituents variables to the final ntuple, also systematic aware.**
- Running on one **LLJ1 MC background** (JZ8 slice) dataset with 20k events, selecting antikt10UFO jets.
 - $p_T > 200$ GeV and $|\eta| < 2$.
 - Max 3 jets kept per event.
 - 1 event = 2.1 kB, about $\sim 73\%$ of size consists of constituents info.



BIGGEST CHALLENGES

➤ **Framework:**

- Trigger studies on jets, requires much more statistics (DxAOD PHYS) → constituents turned off with flag.
- Run on LLJ1 run 3 data, apparently current trigger list not supported by EasyJet for data.
- Inclusion of truth info matching reco jets in the final selected antikt10UFO container.
- Check if EasyJet default jets calibration is correct for our analysis.

➤ **Machine Learning:**

- Definition of new edge weights, based on values before constituents transformation.
- Trial of new edge features.
- Event-level GNN by giving as input both jets per event at the same time instead of single-jet.