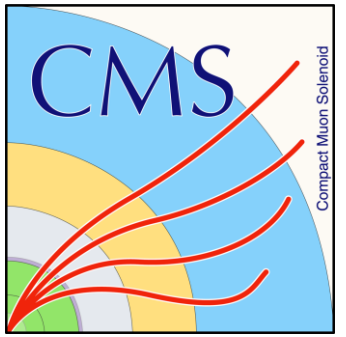


Flavor tagging for Boosted Jets

L.Longo[1], D. Troiano[1,2]

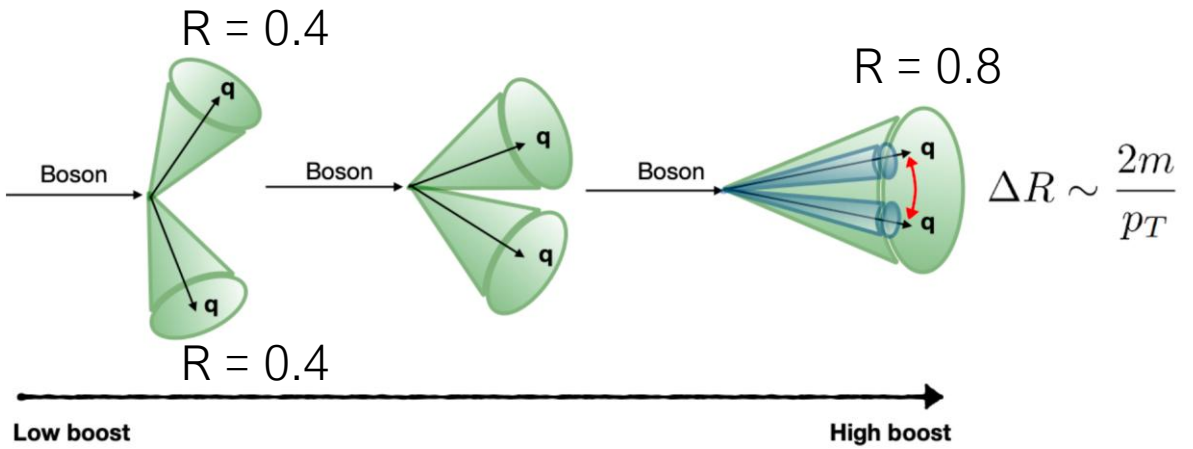
1 INFN of Bari

2 University of Bari

Boosted jets tagging

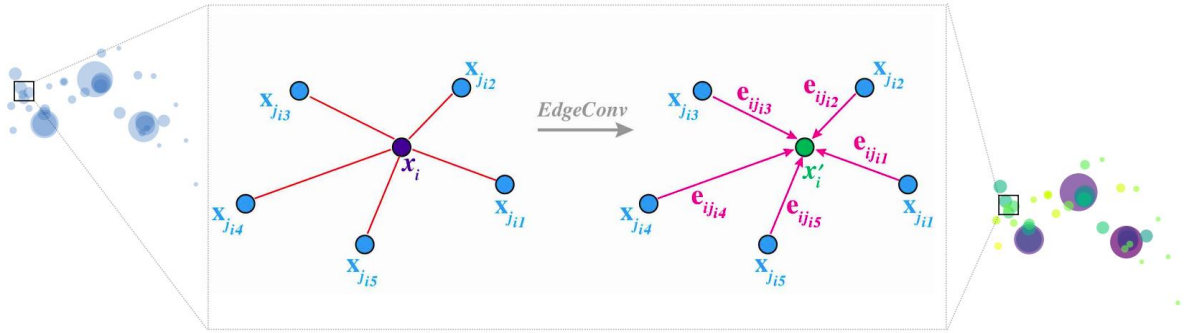
Hadronic decay products of highly boosted particles are collimated.

- Particles merged in one anti-k_T jet (radius R = 0.8 → AK8).
- CMS developed many algorithms to distinguish decaying resonances (*e.g.*, H → $b\bar{b}$) from QCD.



ParticleNet-MD (PNet-MD): state of art for Run 3 boosted jets tagging.

- Jet seen as an unordered set of particles.
- Make a features vector for each jet constituent and secondary vertex based on the features of the *k*-nearest neighbors.



H. Qu, L. Gouskos, PRD 101, 056019 (2020)

My activities:

- First offline PNet-MD validation on Run 3 data.
- PNet-MD training at trigger level for Run 3 data-taking.

First Run 3 offline PNet-MD validation

Validation done on 2022 data events containing highly boosted $Z \rightarrow b\bar{b}$ + jets (CMS-DP-2024-055).

- Events categorized in five PNet-MD score regions.
- Lowest score region almost completely populated by QCD events and not included in the validation.

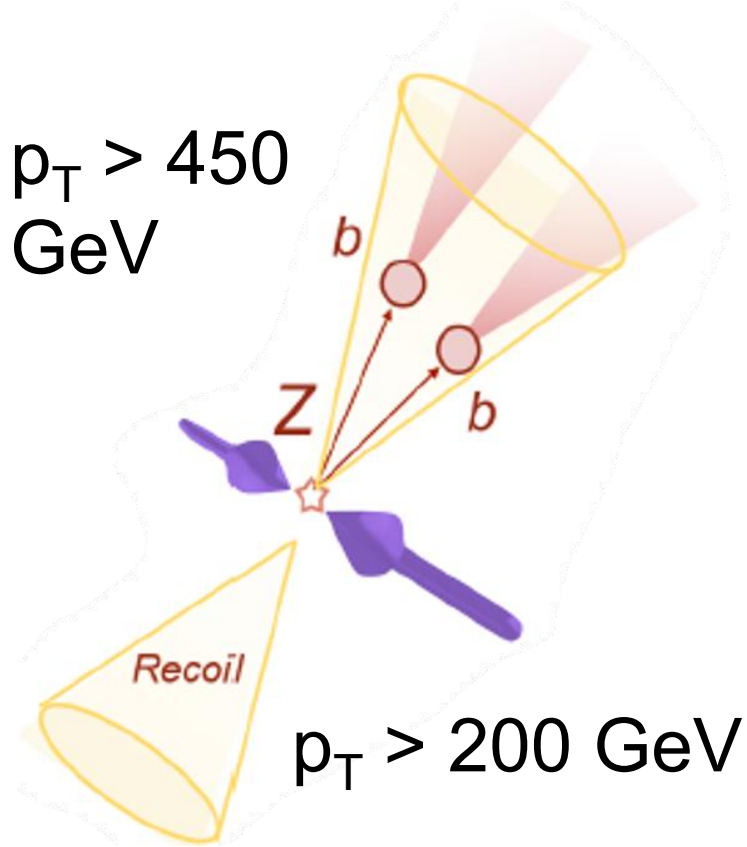
Likelihood fit on the soft-drop mass (m_{SD}) to match data and SM prediction ($Z \rightarrow q\bar{q}$, $W \rightarrow q\bar{q}'$ and QCD).

- m_{SD}: jet mass after removing soft and wide-angle radiation.

Data-driven estimation of QCD multijet background:

- average of the fits of the Z-candidate m_{SD} distributions in nine mass sidebands.

Z-candidate (leading-p_T AK8 jet)

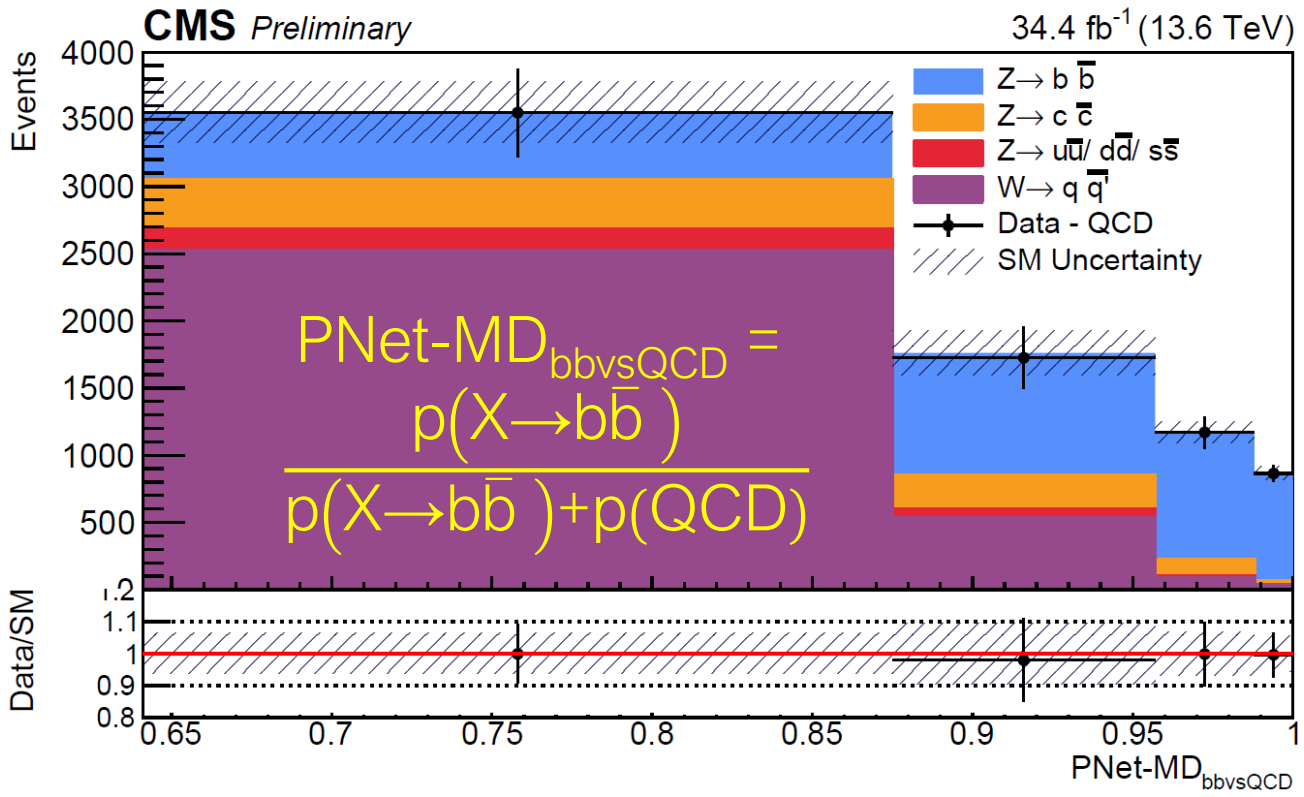


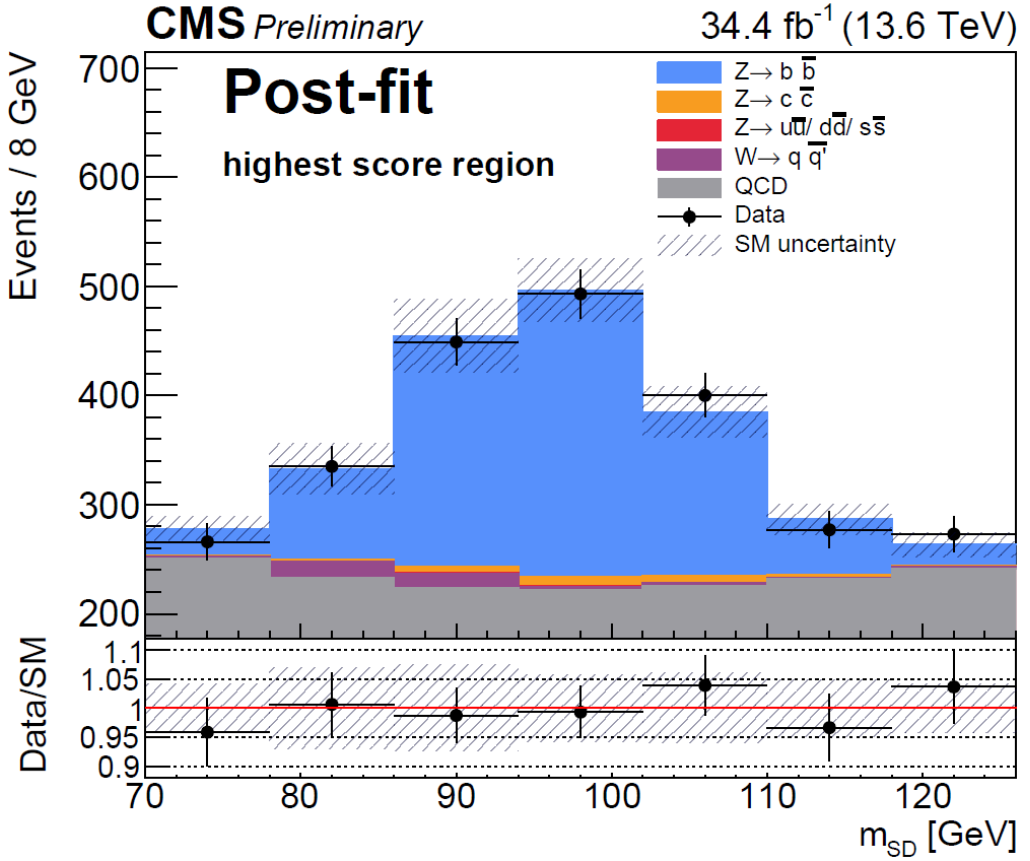
Boosted $Z \to b\bar{b}$ validation results

PNet-$MD_{bbvsQCD}$ offline validation results:

- good data-prediction agreement;
- clear Z-peak in the data distribution.
  - Good signal discrimination from background.

CMS-DP-2024-055





PNet-MD training at HLT

- PNet-MD training at High Level Trigger (HLT) with Winter25 MC simulations.
  - Signal: X→HH (500 GeV < m_X < 6000 GeV, 20 GeV < m_H < 250 GeV).
  - Background: QCD with 4 jets.
  - Goal: to add the newly trained model at HLT.
  - Newly trained model matches the tagging performances of the current one only for p_T > 800 GeV.