

CMS: recent developments in heavy flavor tagging

Clara Ramon on behalf of the CMS
Collaboration

September 30, 2025

HHH Workshop 2025, Dubrovnik
(Croatia)



Introduction

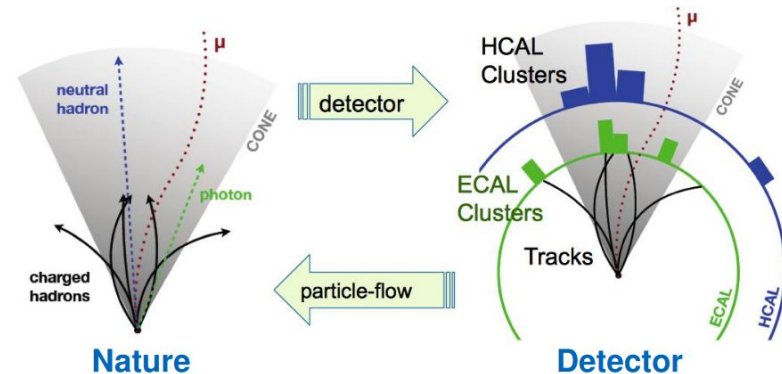
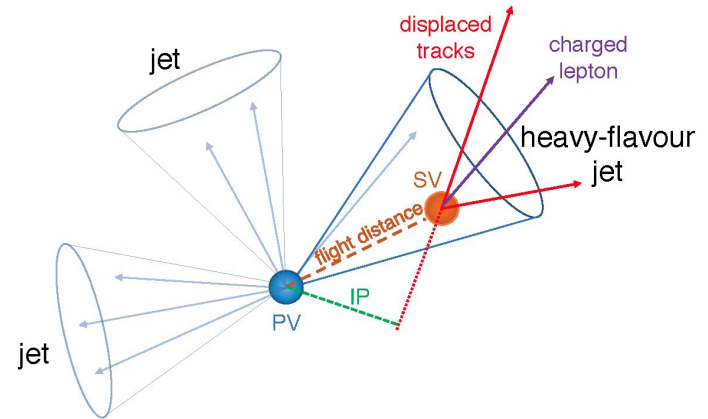
Heavy-flavour tagging allows to identify jets originating from a c- or b-quark by using:

- Jet constituents
- Secondary vertices information

Many standard model (SM) and beyond SM signatures are characterised by heavy flavour quarks in the final state.

- In particular, $H \rightarrow bb$ is the dominant decay channel
- In multi-higgs most sensitive channels target **at least one $H \rightarrow bb$!**

In CMS Particle Flow algorithm is used to define the particles reconstructed correlating information from different detectors.

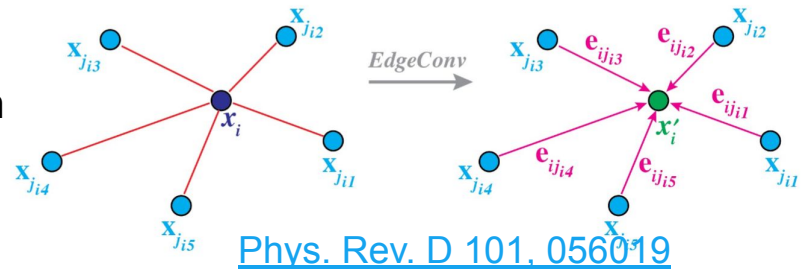
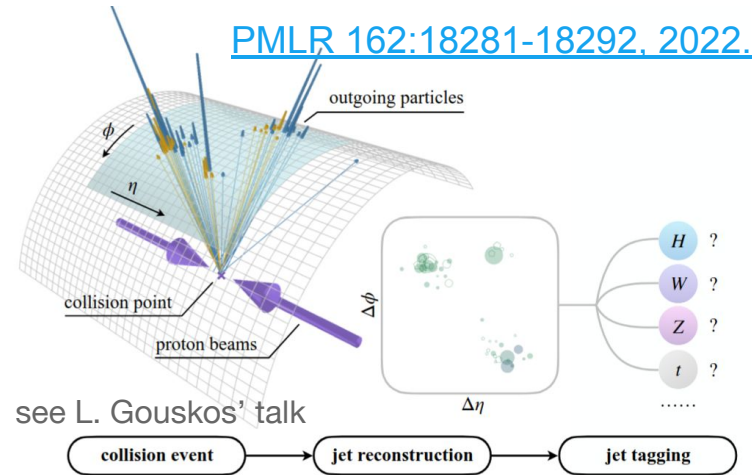


Tagging algorithms in CMS

Large improvement in performance in the last few years by using new Machine Learning techniques.

State of the art in Run 3:

- Offline AK4 tagging based on Particle Transformer:
 - Robust ParT
 - Unified ParT:
includes nodes for s and τ_h tagging
 - Charge Tagger
- AK4 tagging at trigger level based on ParticleNet (DGCNN)



RobustParT

ParticleTransformer (PartT):

- Graph representation by connecting all PF constituents and SVs;
- Multiple interaction features (edges) between PF constituents and SVs:

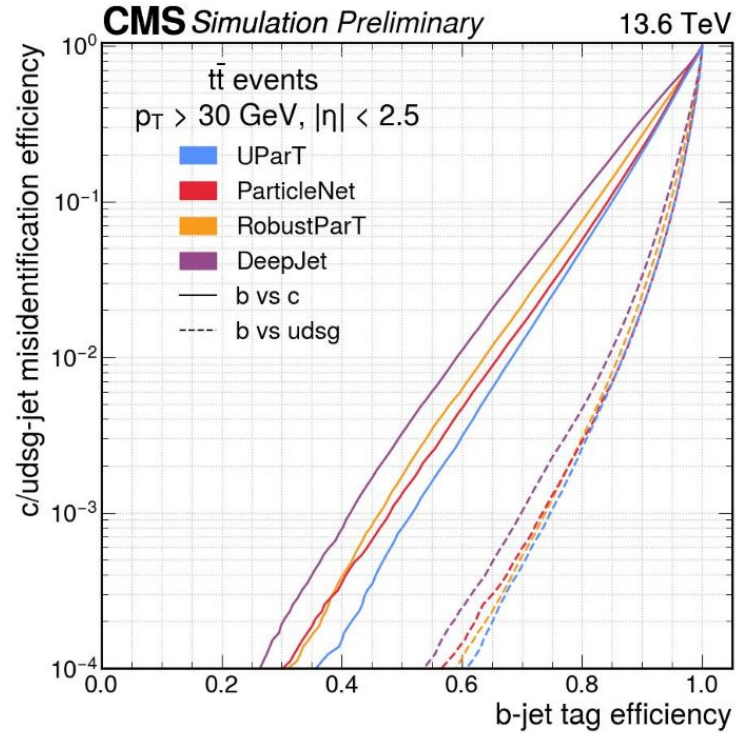
$$\Delta = \sqrt{(y_a - y_b)^2 + (\varphi_a - \varphi_b)^2}$$

$$k_T = \min(p_{T,a}, p_{T,b})\Delta$$

$$z = \min(p_{T,a}, p_{T,b}) / (p_{T,a} + p_{T,b})$$

$$m^2 = (E_a + E_b)^2 - ||\vec{p}_a + \vec{p}_b||^2$$

- Adversarial Training used to enhance the robustness of the model against the mismodeling of our Monte-Carlo (MC) simulation.
- Better comprehension of the jet internal structure.



[CMS-DP-2024-066](#)



UParT

CMS-DP-2024-066

Model

- Model based on Particle Transformer
- Heavy flavour and identification combined, using several tasks:
 - Jet flavour classification;
 - Jet regression;
 - Jet energy resolution.
- Adversarial Training to improve robustness.
- Extra nodes dedicated to s and hadronic tau (τ_h) (split per decay channels) tagging.

$$L = \text{CatEntropy}(x, x_{truth}) + \gamma_{regr} \cdot \log(\cosh(y - y_{truth})) + \gamma_{quantile} \cdot [p_{0.16}(z - z_{truth}) + p_{0.84}(z - z_{truth})]$$

Performance I

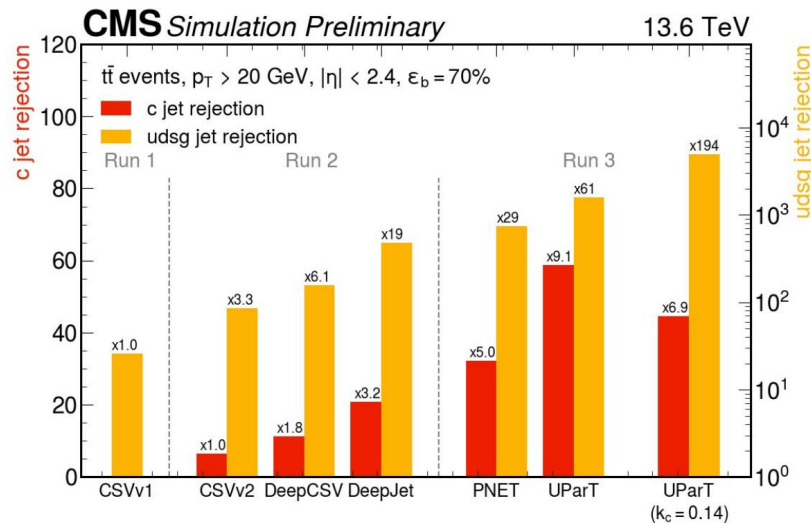
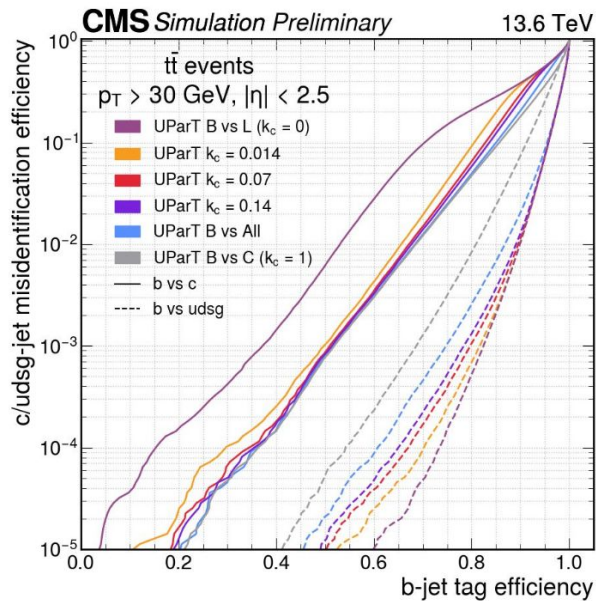
- State of the art b-tagging given by UParT
- Reduces misidentification by a factor of 2 with respect to PNET

$$B \text{ vs All} = \frac{\text{prob}(b)}{\text{prob}(b) + \text{prob}(c) + \text{prob}(uds)}_g$$

$$B \text{ vs L} = \frac{\text{prob}(b)}{\text{prob}(b) + \text{prob}(uds)}_g$$

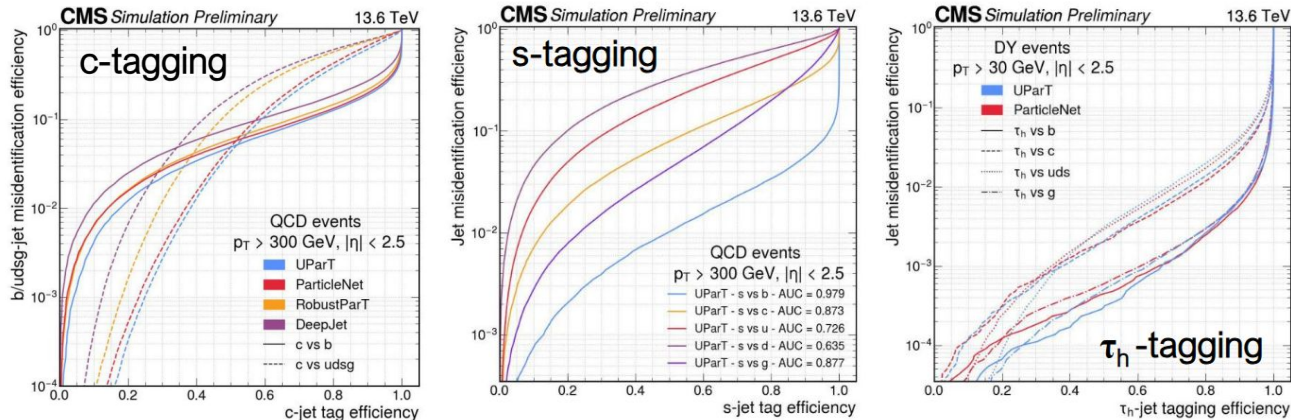
$$B \text{ vs C} = \frac{\text{prob}(b)}{\text{prob}(b) + \text{prob}(c)}$$

$$B \text{ vs All weighted} = \frac{\text{prob}(b)}{k_c \cdot \text{prob}(c) + (1 - k_c) \cdot \text{prob}(uds)}_g$$



Performance II

- UPart reduces c tagging misidentification by a 40% (with respect to PNet)!
- Specific **s-node** for the first time in the CMS experiment!
 - Performances indicates we can achieve a low efficiency s-tagger
- **Tau-tagging:** ParticleNet and UParT show similar performances. ParticleNet performs better at high misidentification rate and UParT at lower rate.





Calibrations

CMS-DP-2024-025

Scale factors on 2022EE

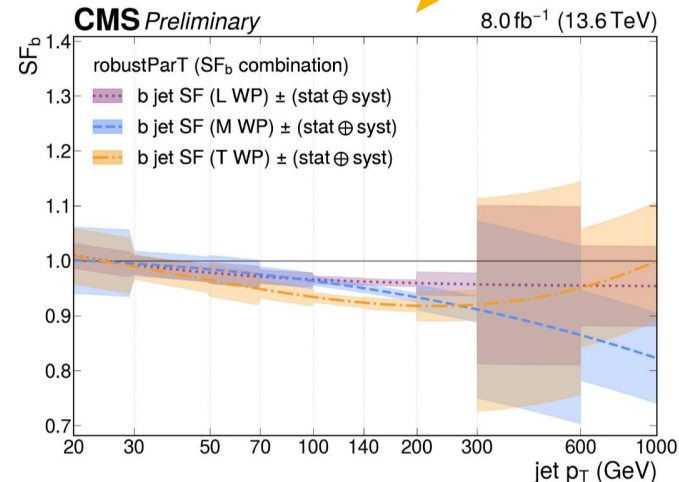
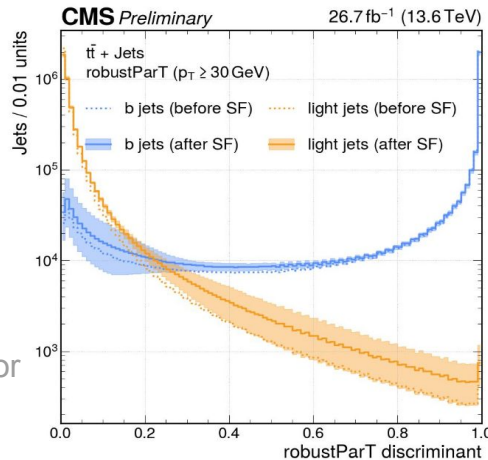
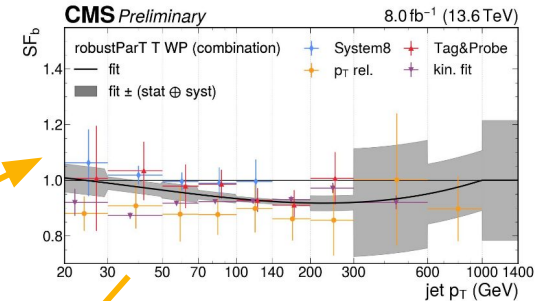
Calibrations for the taggers and data-to-simulation scale factors (SF) derived for b-jets, light-flavoured jets and c-jets.

For b tagging:

- correction in bins of b tagging **working points**
- correction of the full b tagging discriminant **shape**

B tagging efficiencies are computed both on multijet and tt events using different **complementary methods** to ensure robustness of the calibrations.

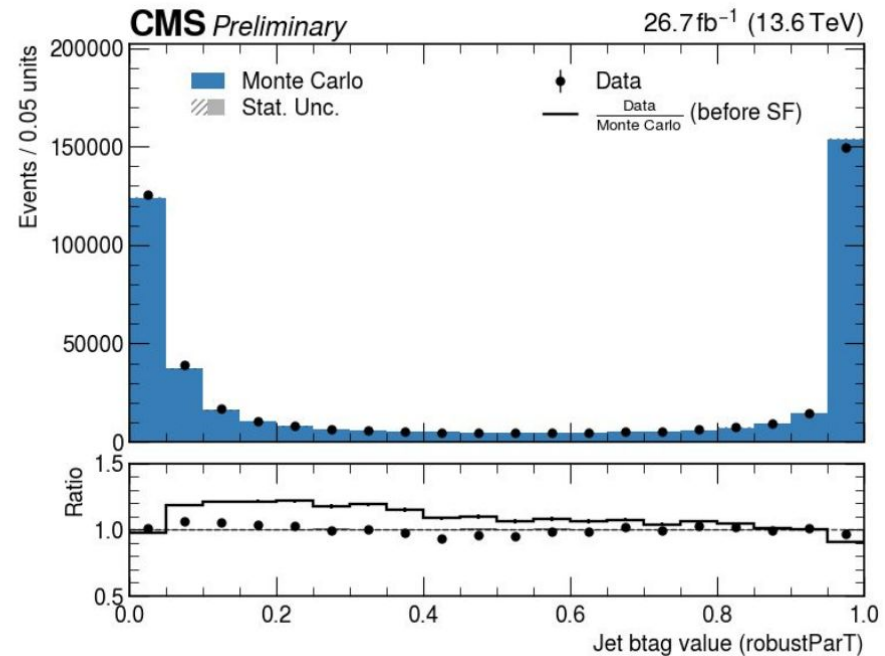
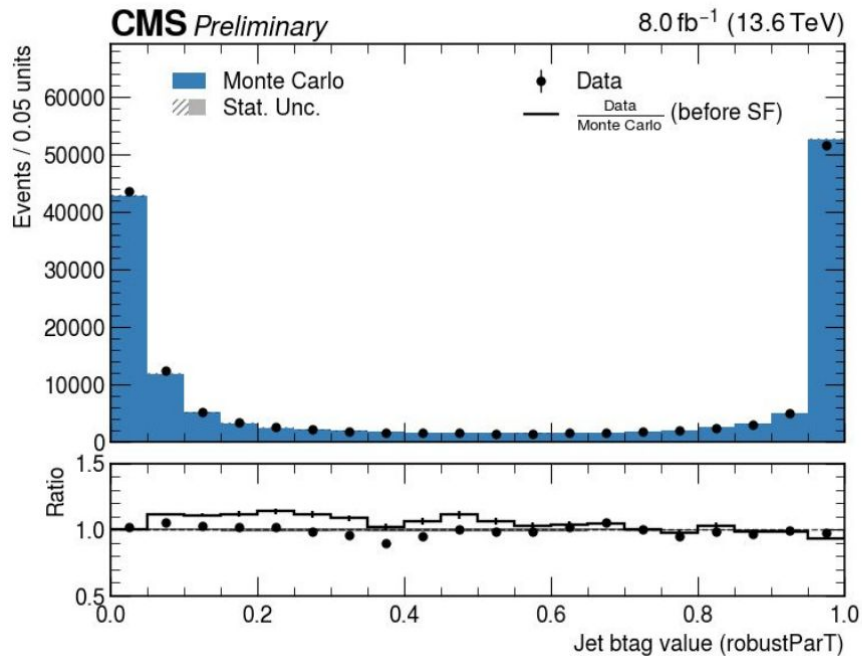
SFs obtained are compatible with 1 in most of the p_T range:

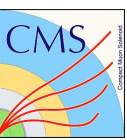


* Showing 2022 as an example, same procedure for other years

Discriminator after calibrations

Good data/MC agreement achieved after calibration in the shape of the discriminator for both eras of 2022 data-taking period:





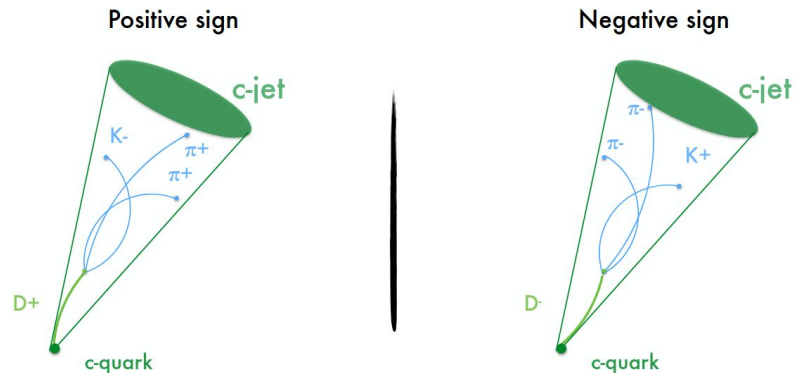
Charge Tagger

CMS-DP-2025



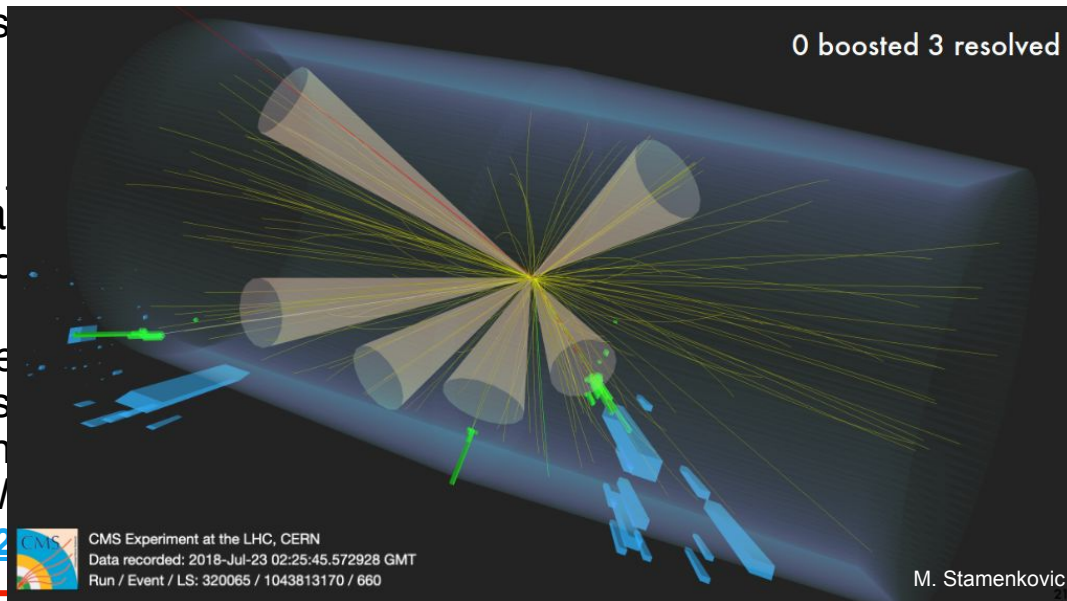
Tagging the charge

- Current flavour tagging algorithms are **charge agnostic**
- Traditional way of assigning the Jet Charge: $Q_\kappa = \frac{\sum_i q_i (p_T^i)^\kappa}{(p_T^{jet})^\kappa}$ where the κ is the regularization parameter of the jet charge. Its value can be tuned to get better separation between positive and negative jets, and we will restrict to two representative cases $\kappa = 0.5$ and 1.
- Using the candidates in the jet we should be able to have some separation between positive and negative charged jets.
 - Recent work in CMS: discriminate hadronically decaying W^+ , W^- , and Z DGCNN based on PNet ([DP2024-044](#))



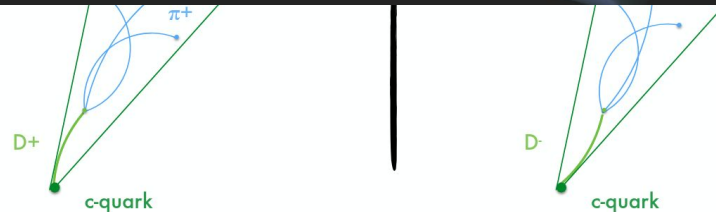
Tagging the charge

- Current flavour tagging algorithms
- Traditional way of assigning the regularization parameter of the separation between positive and negative charged jets
representative cases $\kappa = 0.5$ and
- Using the candidates in the jet we
positive and negative charged jets
 - Recent work in CMS: discriminating
hadronically decaying W^+ , W^-
DGCNN based on PNet ([DP2](#))



Physics interest for Higgs analysis:

- Define signal vs control region based on opposite-sign vs same-sign signatures
- Improve pairing (HH, **HHH**, VH...)



Samples and selection

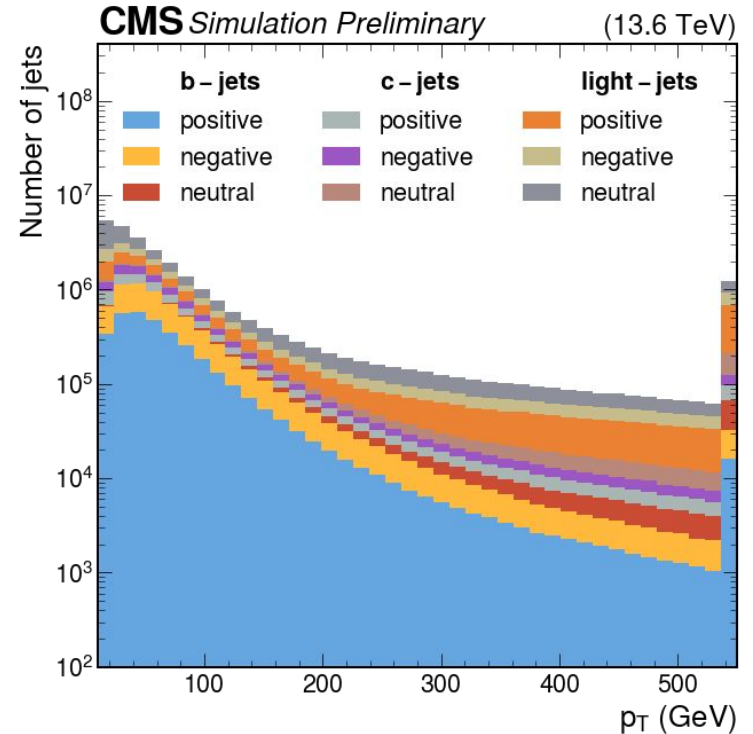
Mixed monte carlo simulation, with a proportion that ensures no flavour class is underrepresented:

- **QCD multijets** (46%)
 - **Top-antitop** ($t\bar{t}$, 15%)
 - **Single Higgs** (ggH (7%), VBF (7%), ZH 21% \rightarrow bb/cc)
 - **$t\bar{t}HH \rightarrow bb$** (4%)
- Provides **heavy-flavour charged jets** (from W , H , top) and **light-flavour jets** (charged uds or neutral gluons).
 - Selected jets are AK4 jets with $p_T > 10$ GeV and η between -2.5 and 2.5 .

Truth labels:

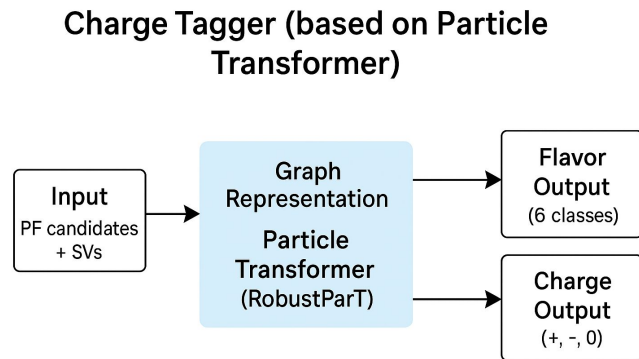
Flavour: For heavy-flavour jets, the topologically matched hadron defines the flavour; for light-flavour jets, it comes from the matched parton.

Charge: Taken from the matched parton. If two $b(c)$ -hadrons are matched, the charge is set to 0. For leptonic decays, the matched lepton defines the charge.



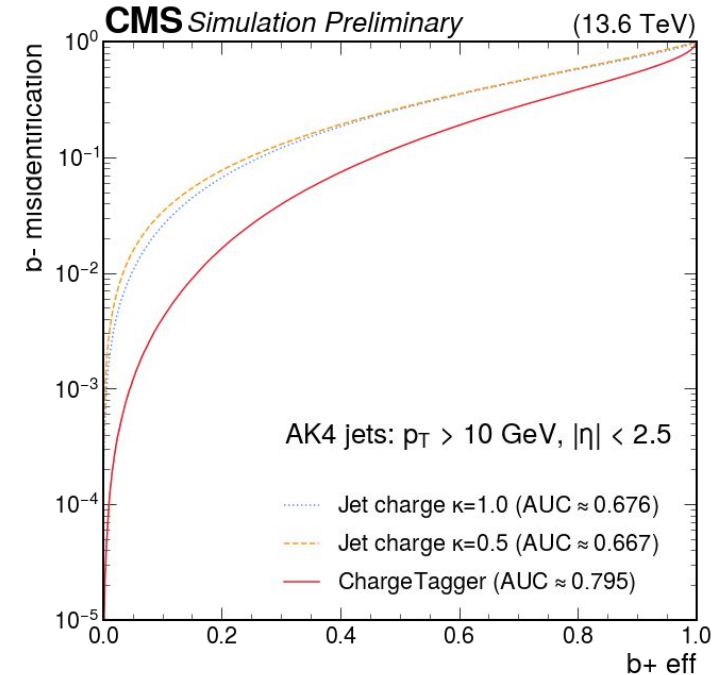
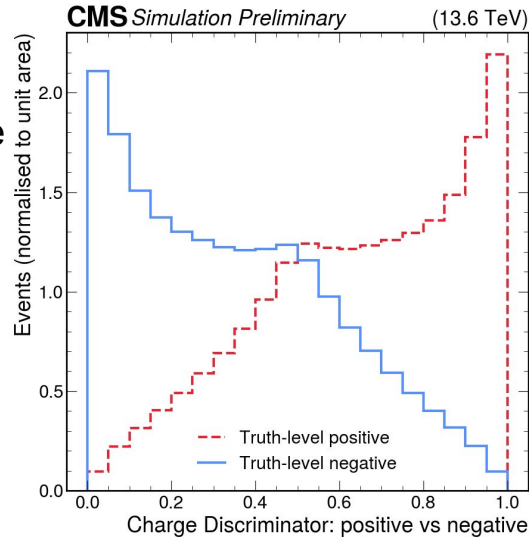
Charge Tagger description

- **Base model:** Particle Transformer (same architecture as RobustParT), graph-based representation of PF candidates (particles) + SVs.
- **Modifications:**
 - Added **charge info** (Charge of the PF charged candidates + SV associated charge).
 - Added **particle ID + tracker hit info** for charged PF candidates.
- **Tasks:**
 - **Flavor classification** (6 classes).
 - **Charge classification** (3 classes: +, -, neutral).
 - **Loss:** sum of task-specific losses.
- **Training:**
 - Implemented in **b-hive framework**.
 - Dataset: 27M jets (train), 5M jets (test).



Binary classification for b

- Take the classifier as binary:
Positive vs Negative
- Negative vs Positive shows the same performance
- Good discrimination between both classes **AUC ≈ 0.8**
- The Charge Tagger shows a **20% better performance** than the traditional cut-based method.



Comparing with LEP:

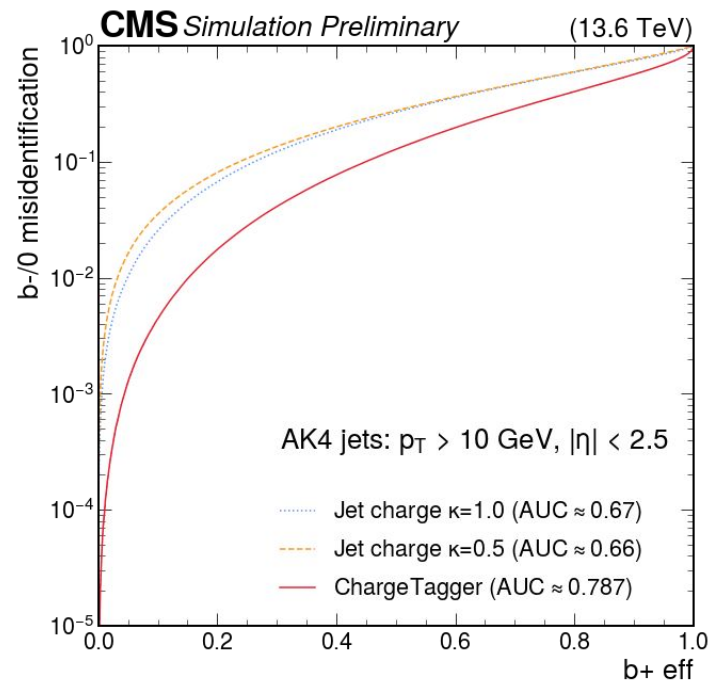
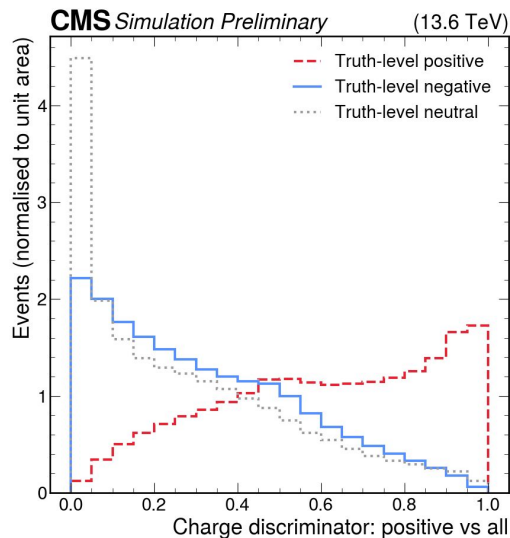
DELPHI NN to distinguish b vs \bar{b} :

- AUC ~ 0.75 [Eur.Phys.J.C40:1-25,2005](https://arxiv.org/abs/hep-ex/9605005)
- We achieve similar performance as in the LEP environment!

$$\text{Positive vs Negative} = \frac{\text{Out}(\text{positive})}{\text{Out}(\text{positive}) + \text{Out}(\text{negative})}$$

Ternary classification for b

- Take the classifier as binary:
Positive vs All
- Negative vs All shows the same performance
- Good discrimination against the two other classes **AUC ≈ 0.8** .
Similar performance as the binary approach.
- The Charge Tagger shows a **20% better performance** than the traditional cut-based method.

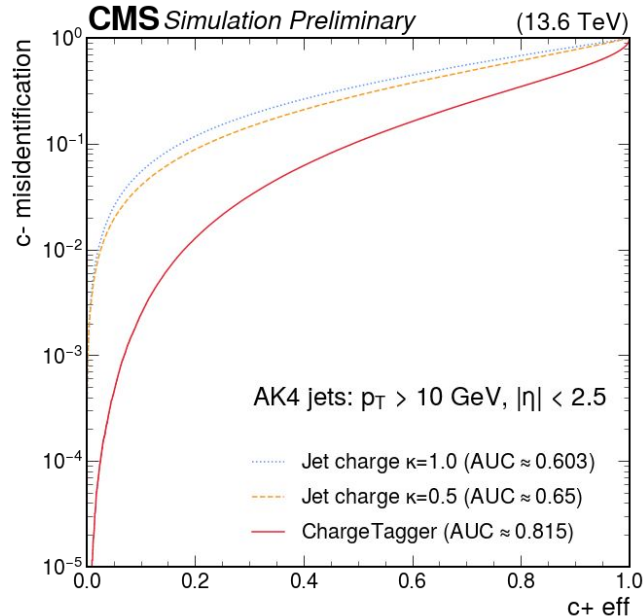


$$\text{Positive vs All} = \frac{\text{Out}(\text{positive})}{\text{Out}(\text{positive}) + \text{Out}(\text{negative}) + \text{Out}(\text{neutral})}$$

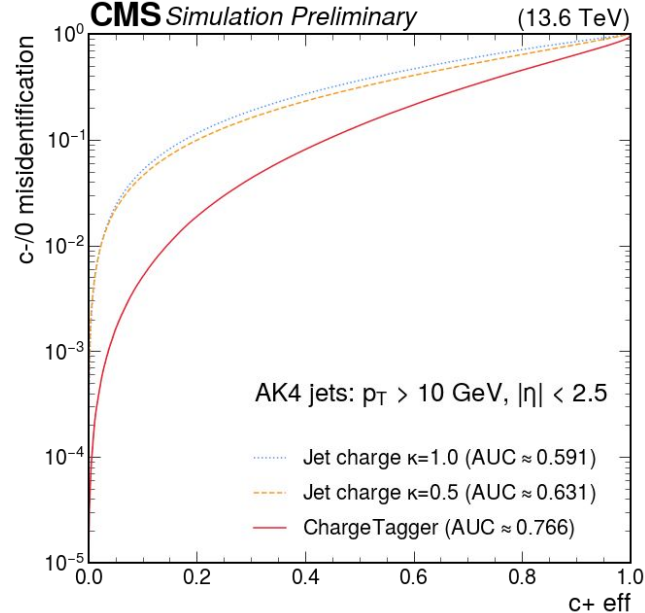
Classification for c

- Binary approach: **better performance than for b**
- Slight larger confusion with neutral class
- Good discrimination against the two other classes **AUC ≈ 0.8** .
- The Charge Tagger shows a **20% better performance** than the traditional cut-based method.

Binary approach:



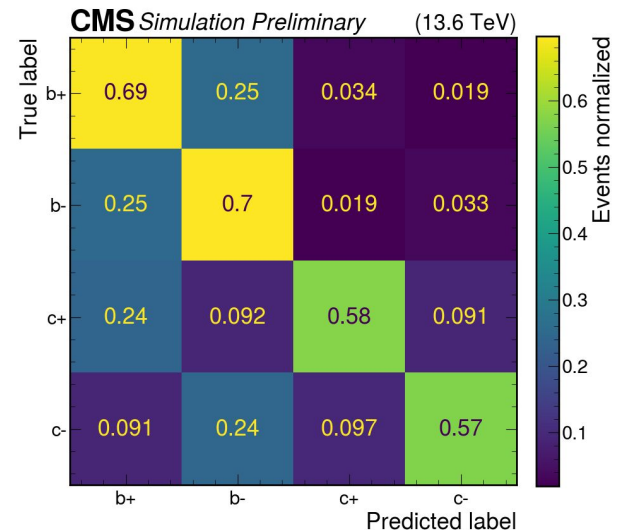
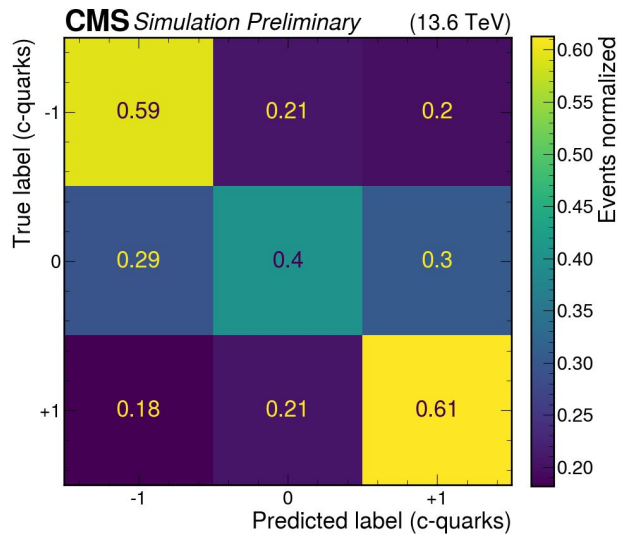
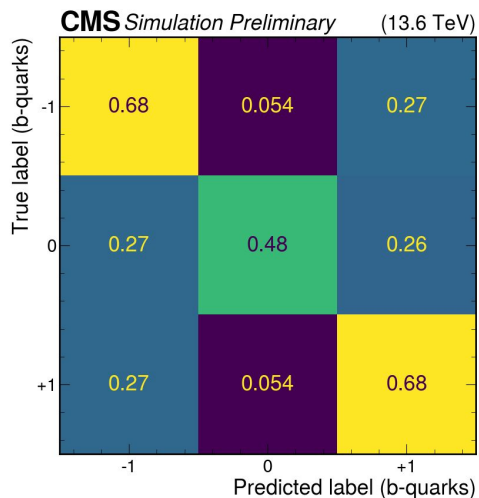
Ternary approach:



Confusion between classes

Using as prediction the **node with the largest score**:

- 70% right assignment for (charged) bs and 60% for cs.
- Larger confusion for cs explained by take into account **confusion between cs and bs**

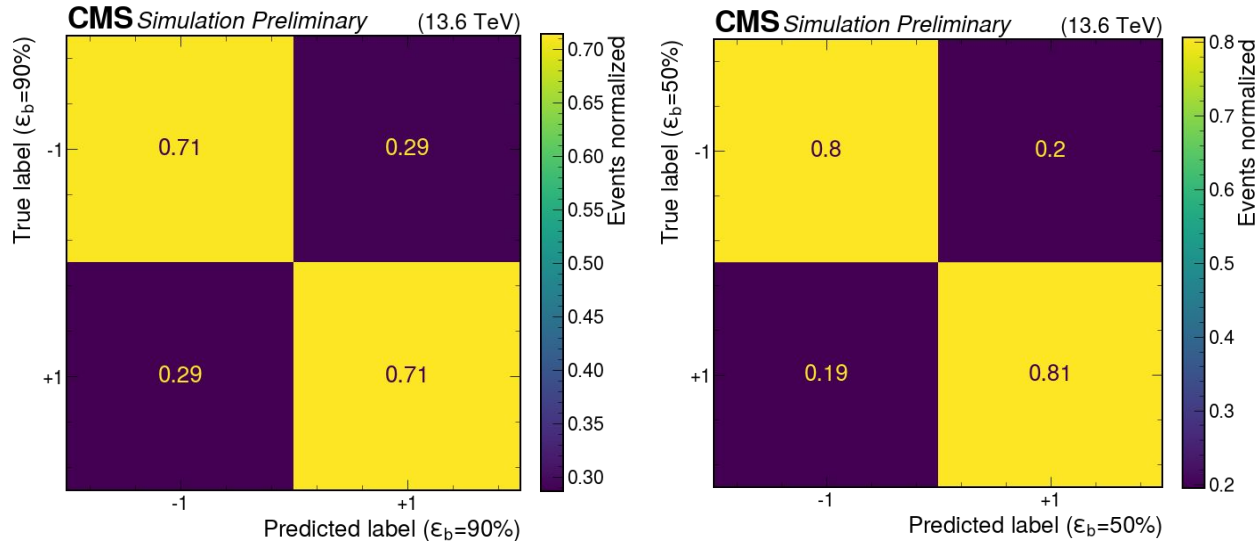


Working point definition for b_s

We can define three preliminary working points for analysis, based on the efficiency of the b^+ (b^- would be symmetric)

Tight selection (50% b^+ efficiency): 80% correct charge assignment

Loose selection (90% b^+ efficiency): 71% correct charge assignment



	Efficiency (%)		
b^+	90	70	50
b^-	55	28	12
c^+	93	73	50
c^-	58	27	10

Summary

State of the art Flavour tagging in CMS takes advantage of graph representation of jets:

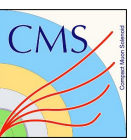
- ParticleNet and ParticleTransformer are used
- These technique **improve background rejection rates** with respect to run 2 algorithms!
 - b-tagging misidentification reduced by a factor ~ 2

Current taggers are able to tag not only b and c-jets but also s-jets and τ_h in a unified model

ParticleTransformer is used for the first time on CMS to **tag the charge of heavy-flavoured jets**

- 20% better performance than the traditional cut-based approach
- Similar performance as the charge tagging for b-jets achieved in DELPHI

*Thank
you!*



Back up

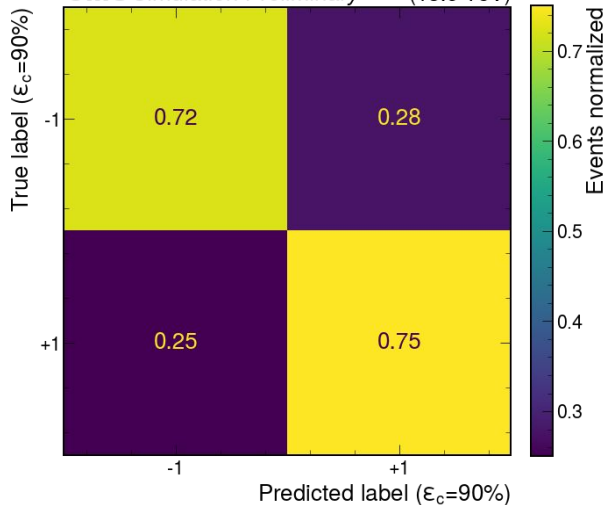
Working point definition for cs

Very similar picture as for cs, slightly better performance, as expected.

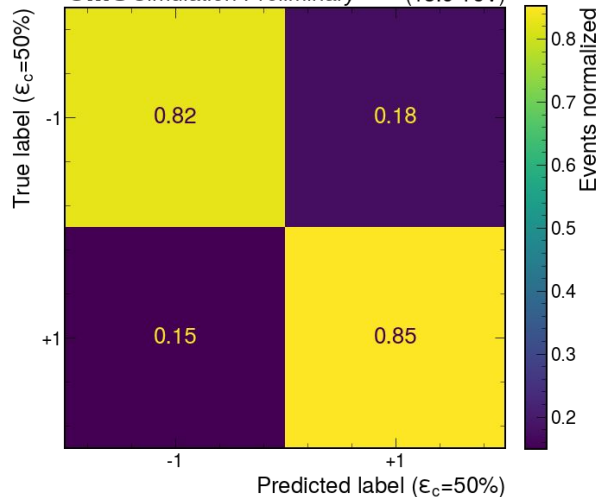
Tight selection (50% b⁺ efficiency): 82% correct charge assignment

Loose selection (90% b⁺ efficiency): 72% correct charge assignment

CMS Simulation Preliminary (13.6 TeV)



CMS Simulation Preliminary (13.6 TeV)

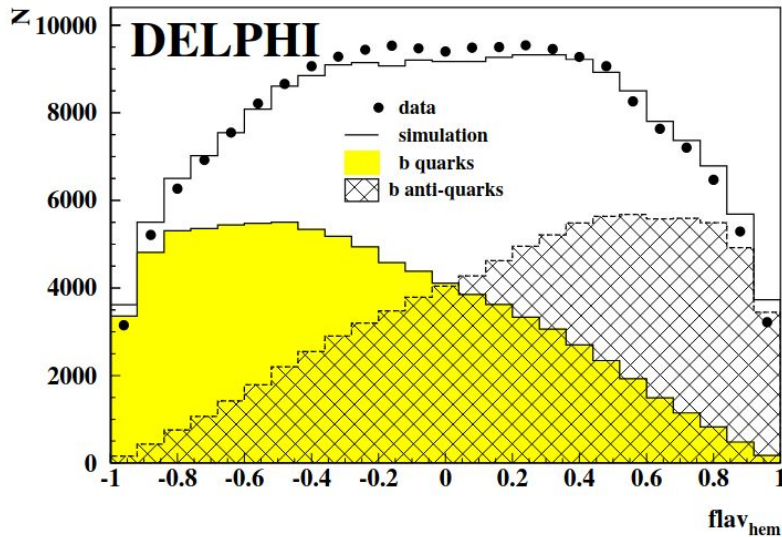


Efficiency (%)

	90%	50%	90%
b ⁺	90	70	50
b ⁻	55	28	12
c ⁺	93	73	50
c ⁻	58	27	10

Comparison with LEP

At DELPHI NN were used to discriminate b quarks and b anti-quarks. Used inputs are the Jet Charge, SV charge, and information about the hadronization and fragmentation.

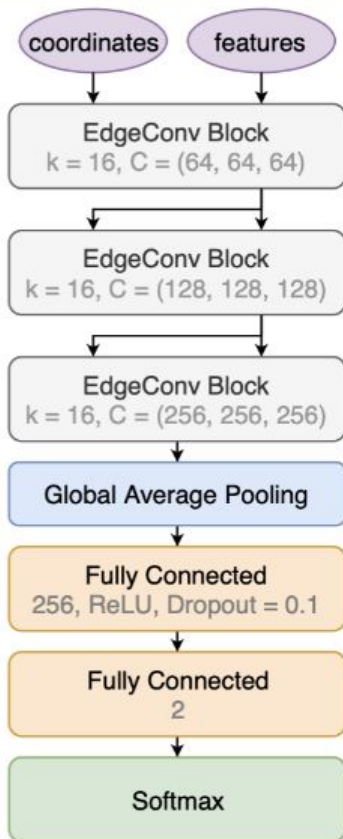
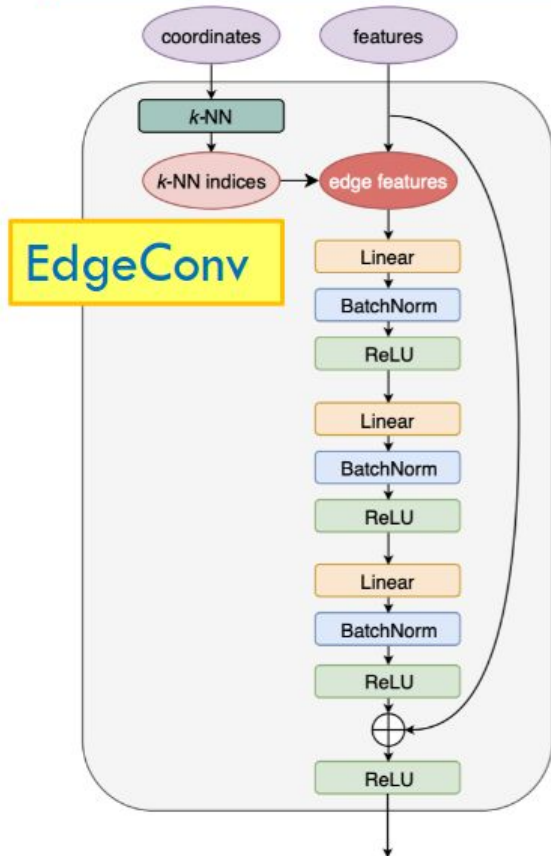


⇒ We can infer the from this plot: AUC ~ 0.75

Our Charge Tagger is performing at the same level as could be achieved in the LEP environment!

[Eur.Phys.J.C40:1-25,2005](#)

ParticleNet architecture



Binary outputs:

$$\text{PNet-MD}_{\text{bbvsQCD}} = \frac{p(X \rightarrow b\bar{b})}{p(X \rightarrow b\bar{b}) + p(\text{QCD})}$$

$$\text{PNet-MD}_{\text{ccvsQCD}} = \frac{p(X \rightarrow c\bar{c})}{p(X \rightarrow c\bar{c}) + p(\text{QCD})}$$

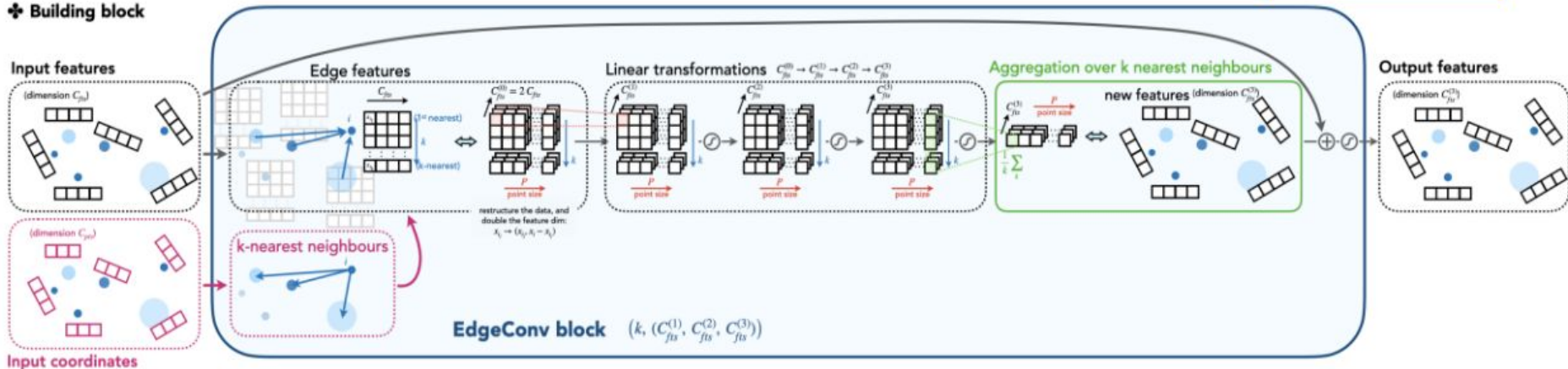
$$\text{PNet-MD}_{\text{qqvsQCD}} = \frac{p(X \rightarrow q\bar{q})}{p(X \rightarrow q\bar{q}) + p(\text{QCD})}$$

$$\text{PNet-MD}_{\text{ttvsQCD}} = \frac{p(X \rightarrow t\bar{t})}{p(X \rightarrow t\bar{t}) + p(\text{QCD})}$$

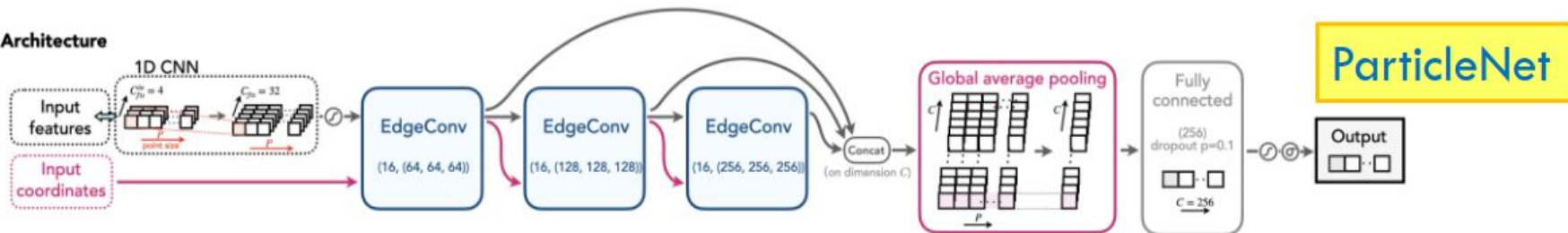
ParticleNet architecture

EdgeConv

Building block



Architecture



WP-based scale factor measurements

b tagging efficiency measurements in multijet events:

- p_T^{rel} – Method to measure the b tagging efficiency in multijet events based on the transverse momenta of muons within jets with respect to the jet axis [13].
- **System8** – Method to measure the b tagging efficiency in multijet events with muons within jets by solving a system of 8 equations [13].
- **Lifetime (LT)** – Method to measure the b tagging efficiency in multijet events with muons within jets based on template fits to the distributions of the JP discriminator and secondary vertex mass [13].

b tagging efficiency measurements in top pair events:

- **tt** – Top quark antiquark pair production.
- **Kin. fit** – Method to measure the b tagging efficiency in $t\bar{t}$ events with two leptons, based on a template fit to an MVA discriminator combining several kinematic observables [13].
- **Tag&Probe** – Method to measure the b tagging efficiency in $t\bar{t}$ events with one lepton, based on a tag and probe method, where the medium WP is applied to one of the two b-jets expected from top pair production, while the other b jet is used as a probe [13].

Combination of b tagging efficiency measurements:

- Combination of b-tagging efficiency measurements from $t\bar{t}$ events (Kin. fit, Tag&Probe) and multijet events with a muon ($p_T^{\text{rel.}}$, System8, LT) using the BLUE method [14], in which the weighted average is determined taking into account the correlation between the systematic uncertainties.

Light jet mistagging efficiency measurement in multijet events:

- **NegativeTag** – Method to measure the light jet mistagging efficiency in inclusive multijet events, using the tracks with negative sign of the impact parameter [13].

ParticleTransformer

[CMS-DP-2024-066](#)

*UParT reduces
PNet c-tagging
misidentification by
~ 40% !*

