Jets and MET in CMS

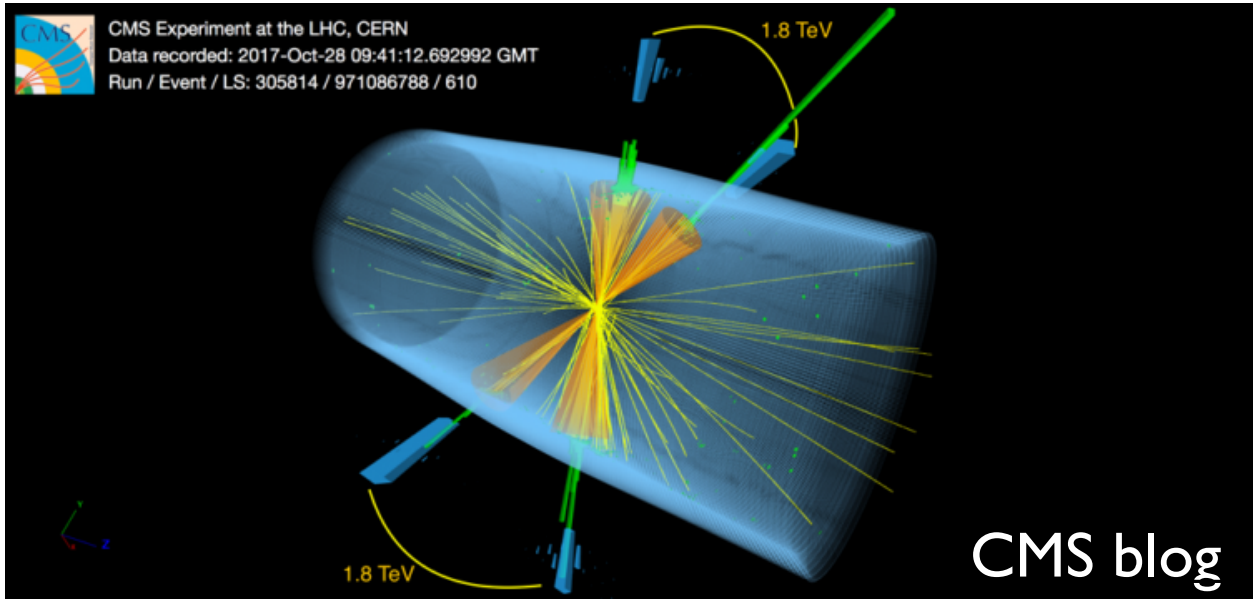
JetMET workshop in Brussels, 2023

Mikko Voutilainen (Helsinki), Alexis Kalogeropoulos (Kansas State)

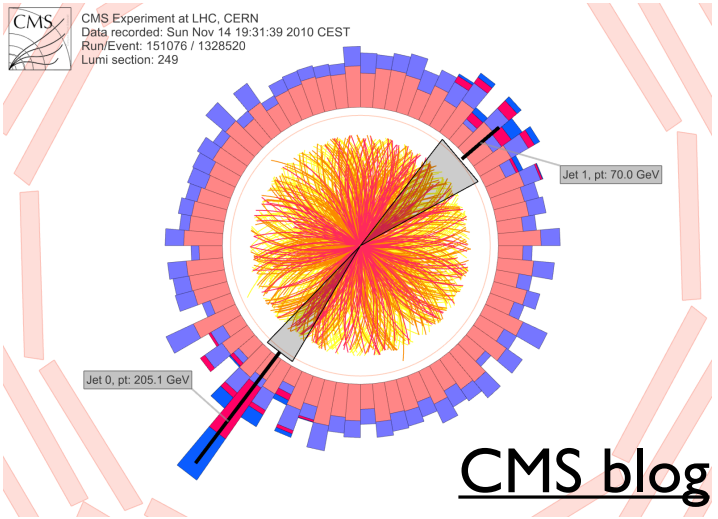
Overview

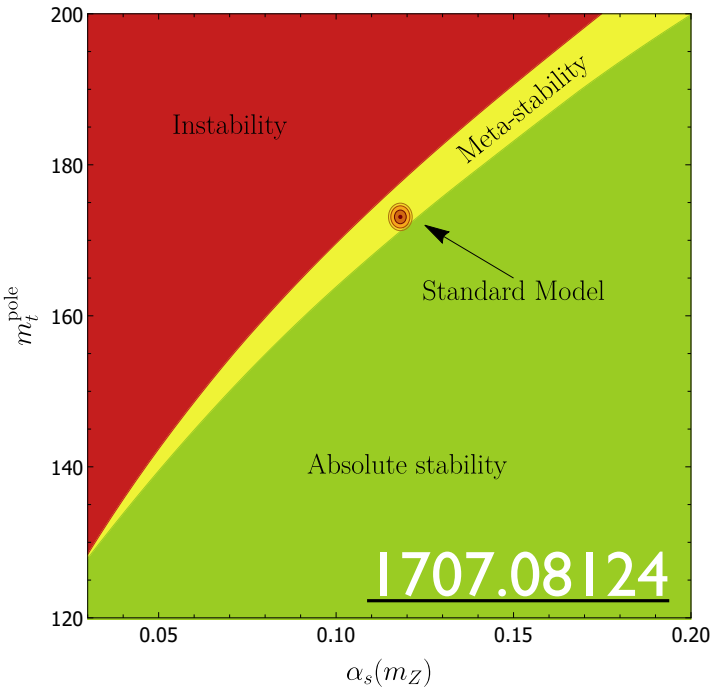
- Introduction
  - jet is the answer, what is the question?
  - building from the bottom to the top, and back
  - what are little jets made of?
- PUPPI core
  - the river threading through JetMET
- Trigger
  - no data, no gain
- Data quality
  - not - a - crap in, crap out experiment
- Jet energy and resolution
  - precision is the game
- Missing momentum
  - what happens in CMS, stays in MET
- JetMET algorithms and reconstruction
  - sublime performance
- Summary and outlook

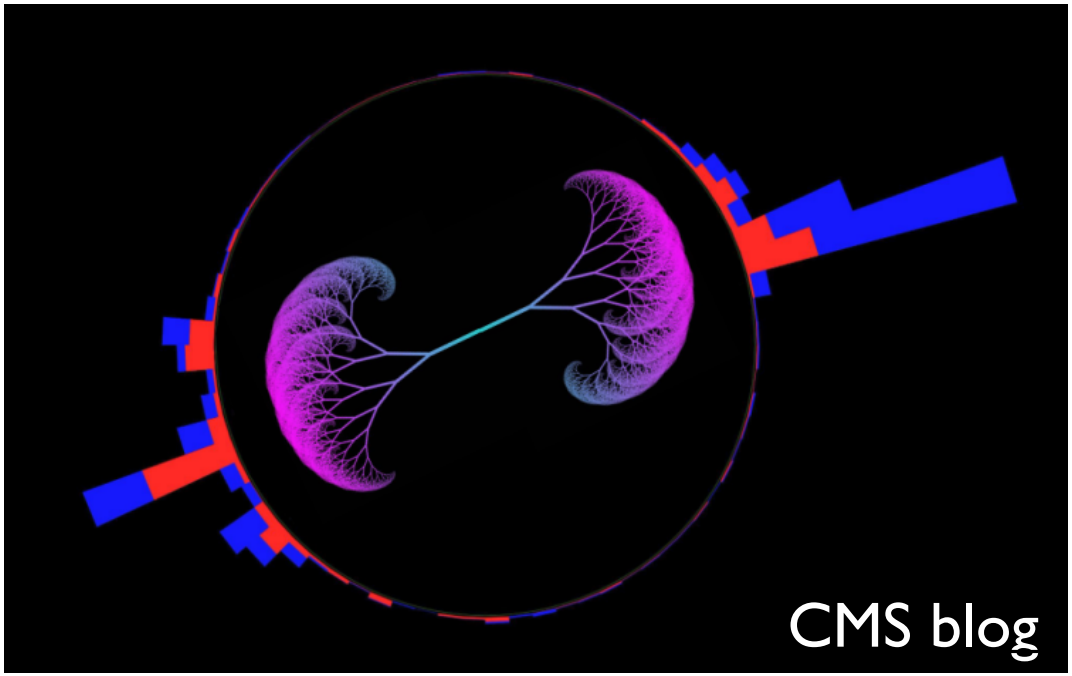
Jet is the answer...



- …what is the question?

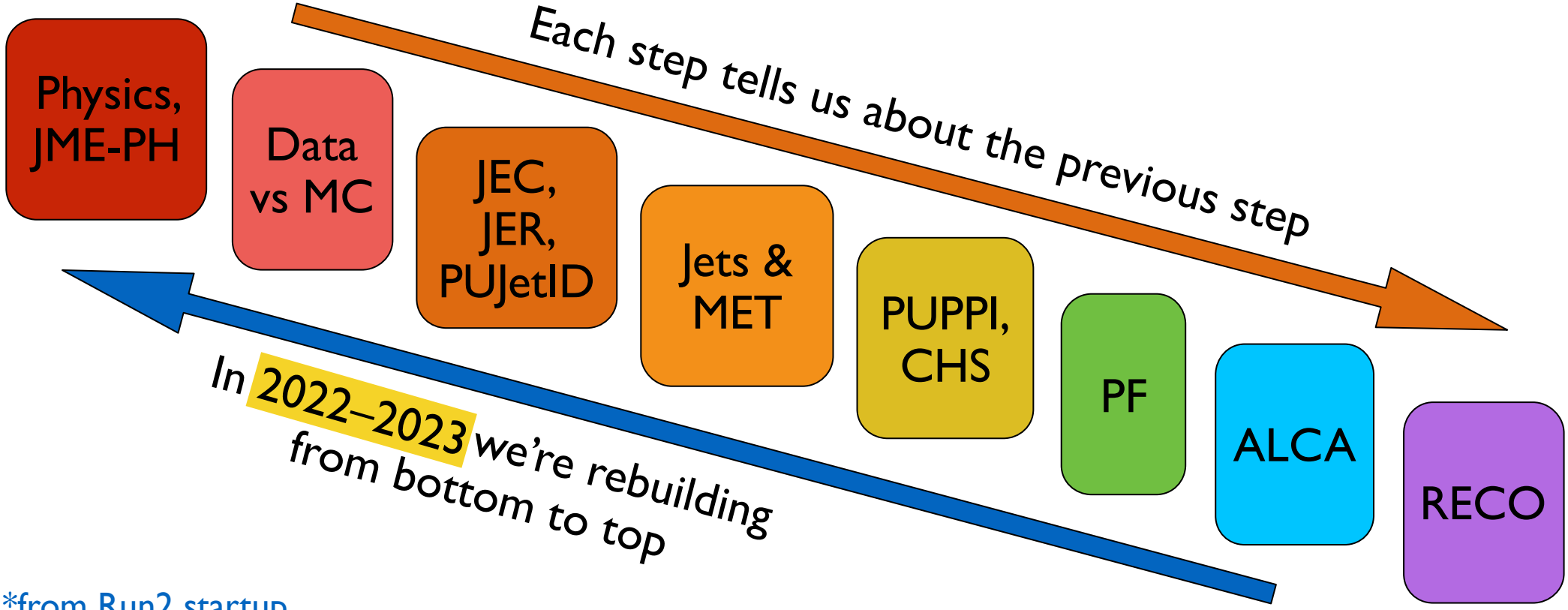






From bottom and back

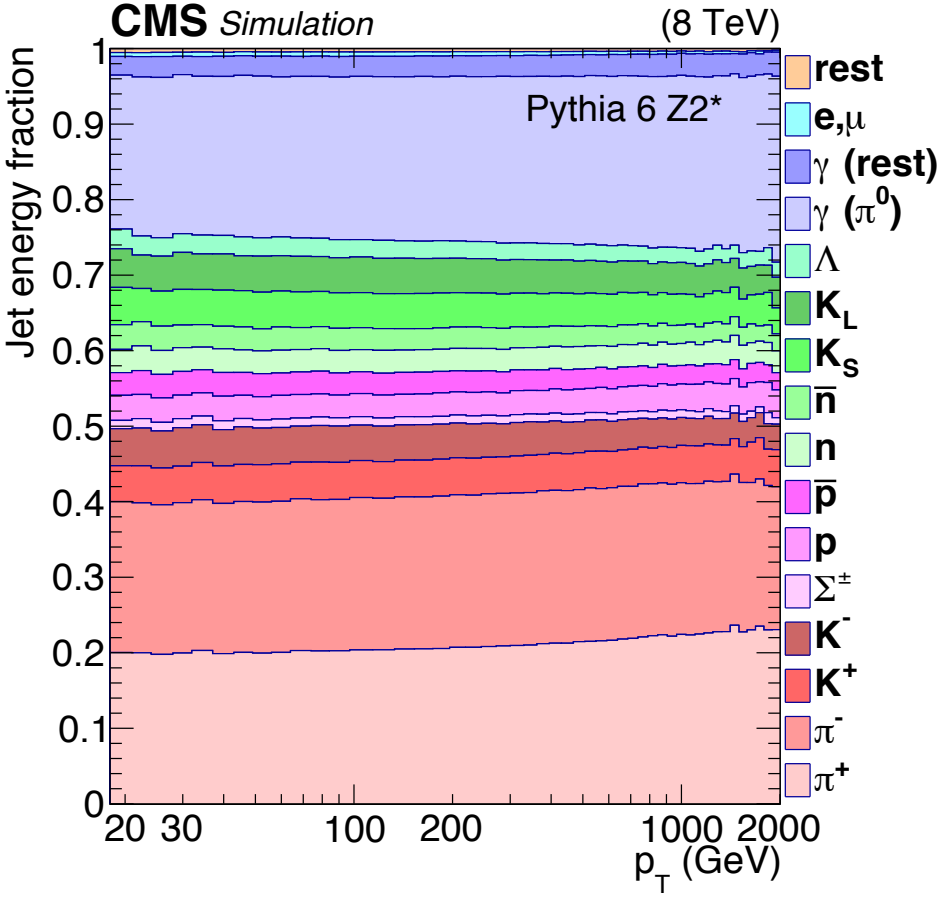
- JetMET is at the end of a long chain of corrections, each of which needs to consistent
- Often subtle problems are only seen at higher levels with more statistics
- We are gatekeepers to physics, and characterise the final performance of our objects



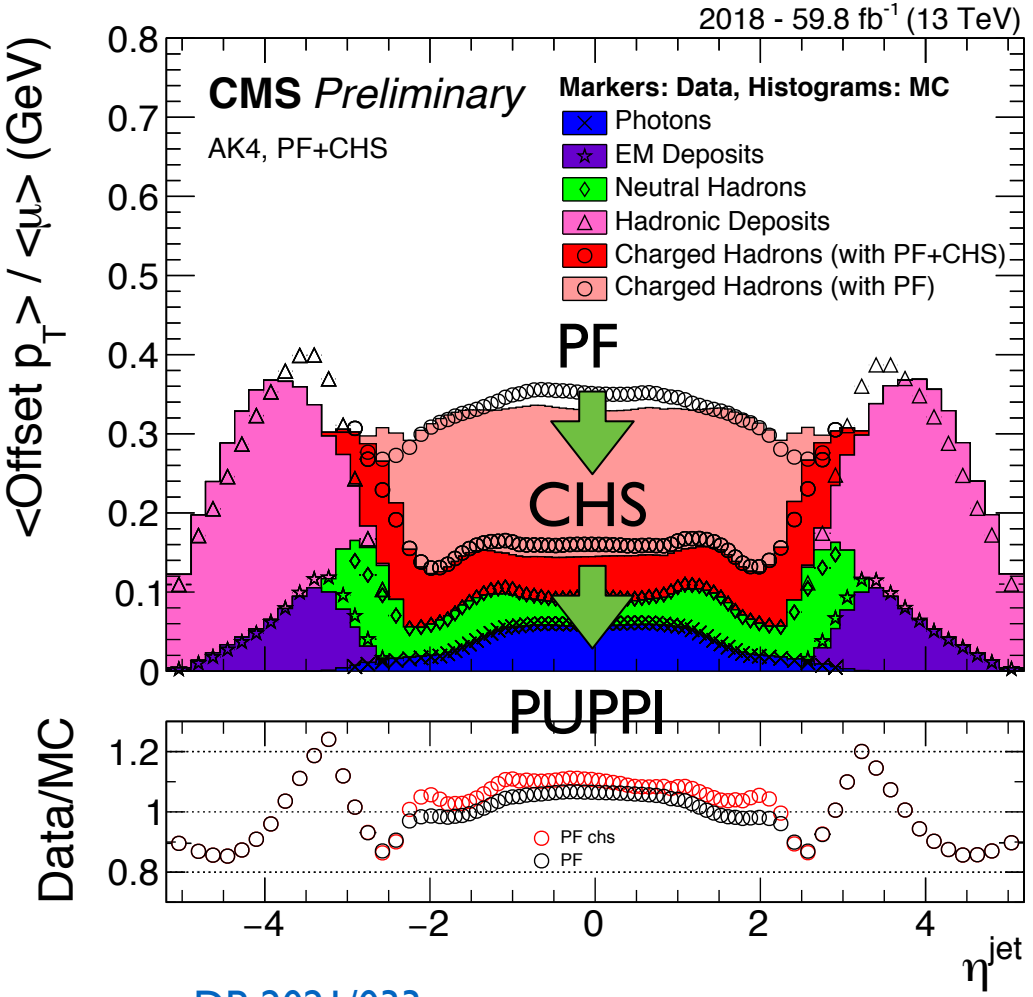
*from Run2 startup

Little jets are made of

- many many particles ($\pi^0 \rightarrow \gamma\gamma$, π^+, π^-, K^0_L, n, p etc.) from many many vertices
- Particle Flow (**PF**) reconstructs 3 types: charged hadrons, neutral hadrons, photons
- Charge Hadron Subtraction (**CHS**) and Pile Up Per Particle Id (**PUPPI**) remove pileup



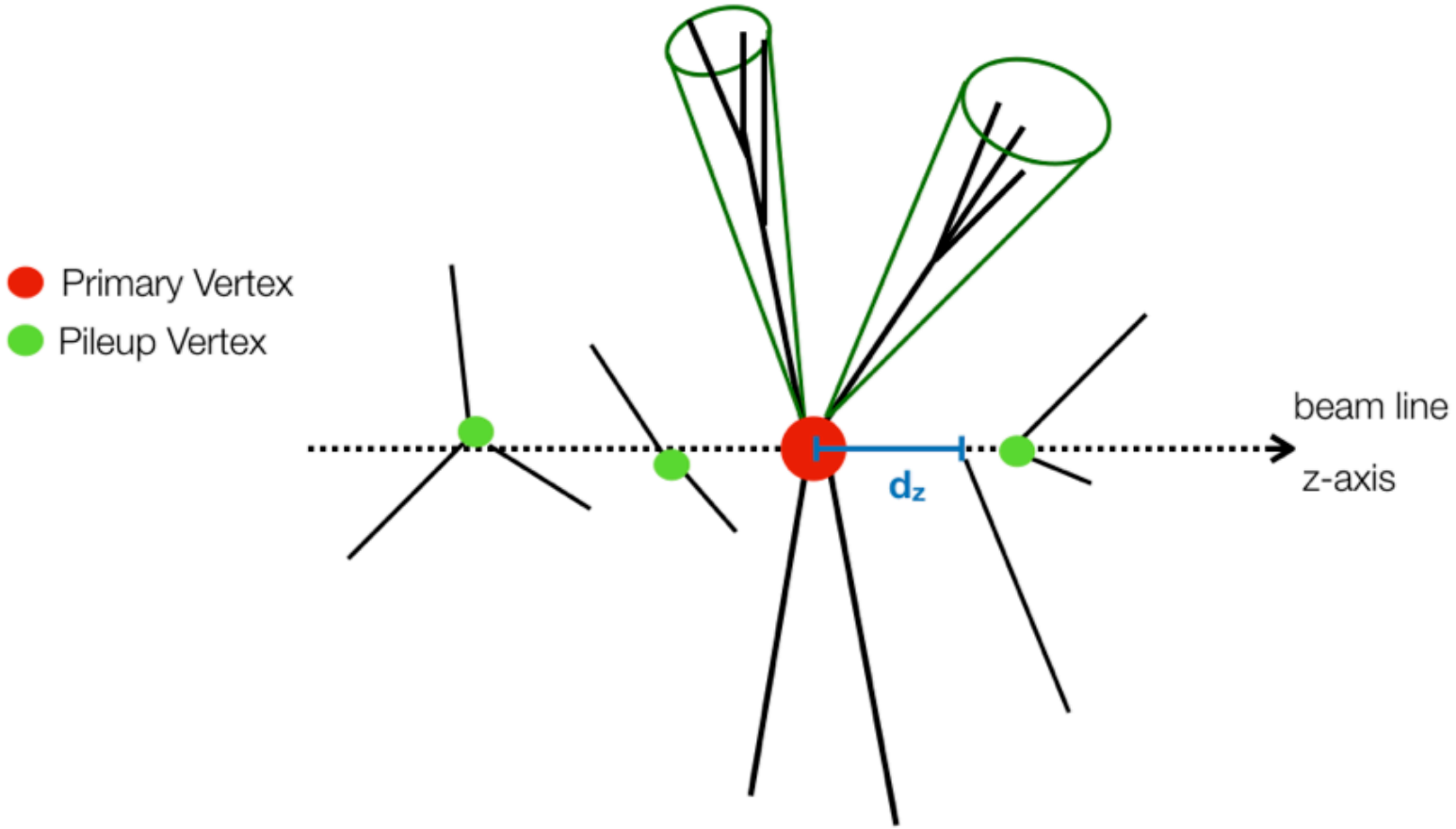
JME-13-004



DP-2021/033

Tracking in PUPPI

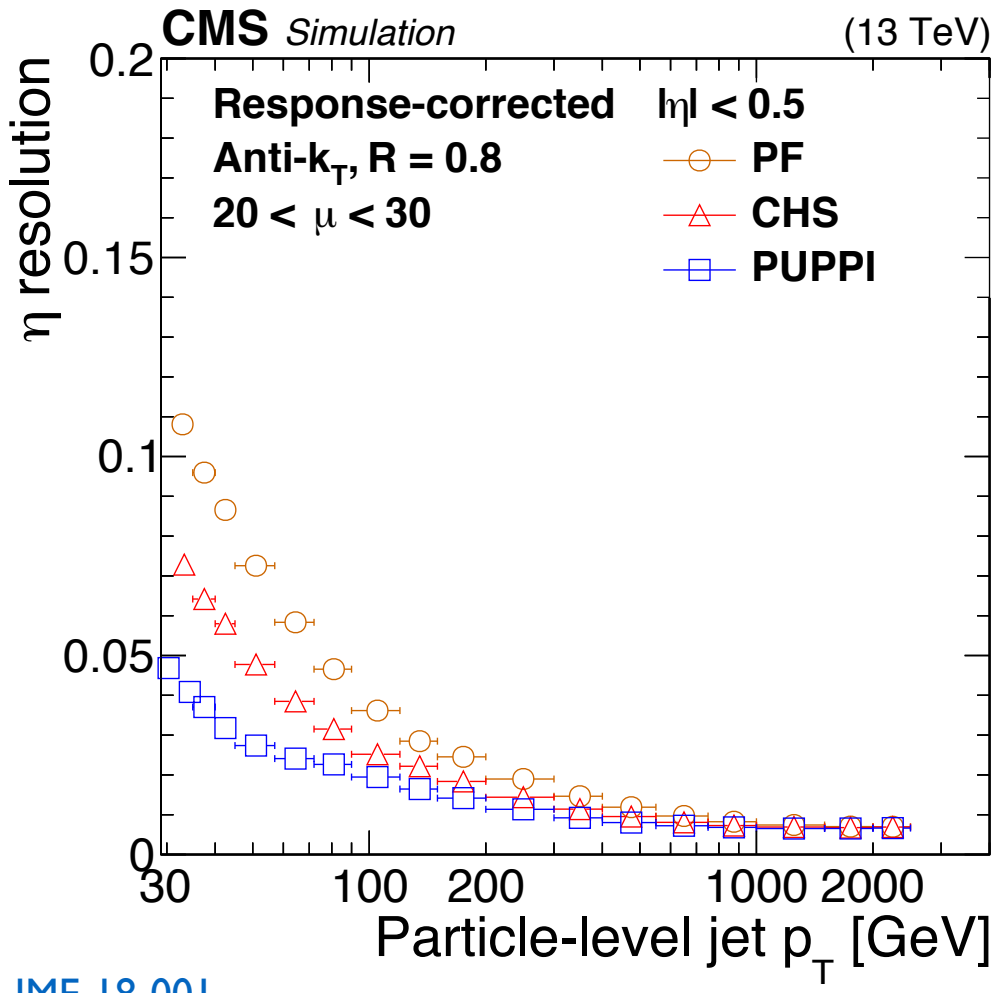
- PUPPI uses 3 particle associations: (1) Pileup vertex, (2) Primary vertex, (3) Ambiguous
  - Pileup vertices give expected ($(p_T/\Delta R)^2$ -> α) distribution for pileup particles wrt primary vertex
  - Ambiguous (e.g. neutral) particles are compared to primary vertex and pileup vertex particles
  - Each particle is given weight from probability of it not being pileup (#1=0, #2=1, #3=w, 0<w<1)
- Track reconstruction and track-vertex association is at the very core of PUPPI
  - also true for PF+CHS in Run 2



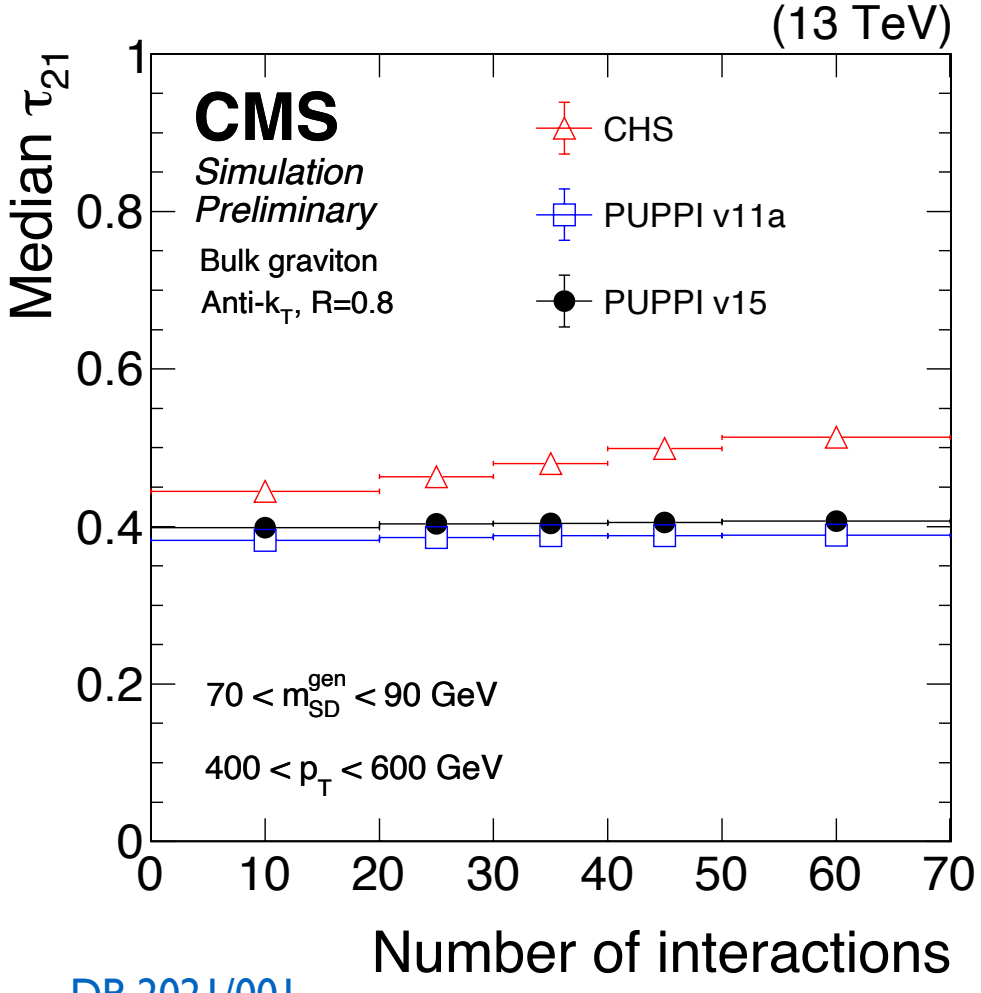
DP-2021/001

PUPPI performance

- PUPPI improves jet and MET resolutions at N_{PU}>40 and even lower for large R
- Just as important is substructure: mass scale, mass resolution, N-subjettiness
- With PUPPI, can perform very detailed jet substructure studies even at high PU



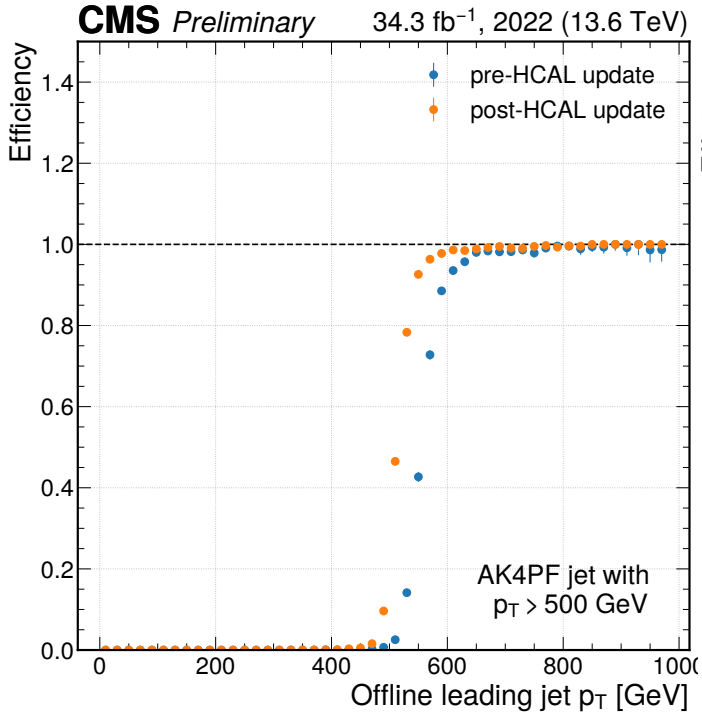
JME-18-001

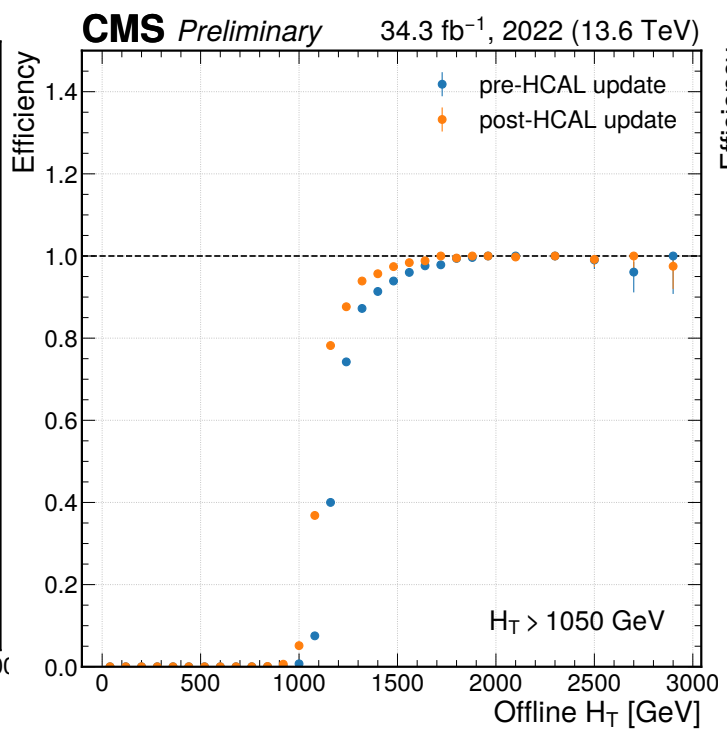


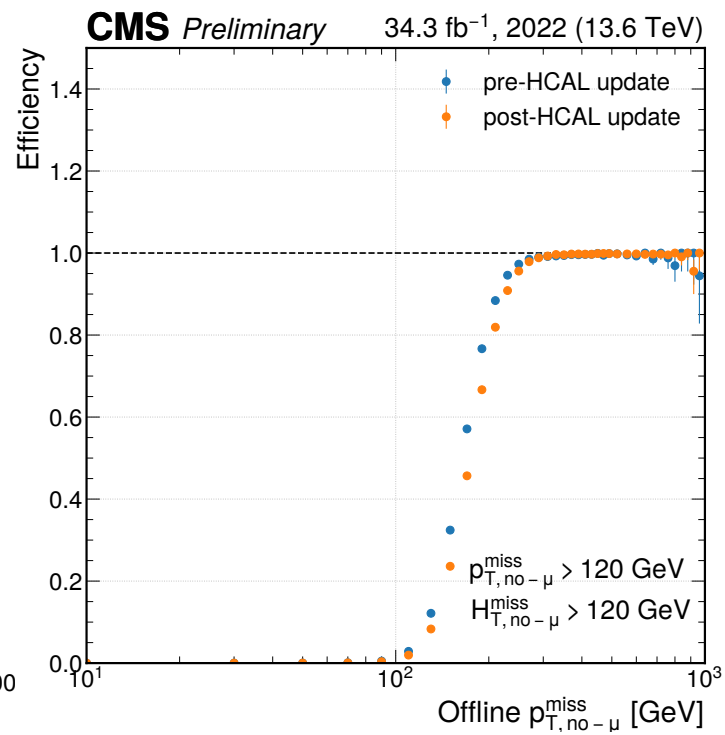
DP-2021/001

Triggers

- Jet and MET data triggered with three main types:
  - Jet triggers with **single p_T** threshold (single jet, forward jet, dijet average; AK4 and AK8 sizes)
  - H_T, i.e. **scalar-sum jet p_T** triggers
  - MET or $p_{T,miss}$ triggers for (negative) **vector-sum jet p_T**
- Triggers are in the front lines to collect data for validation and calibration
  - performing well for Run 3





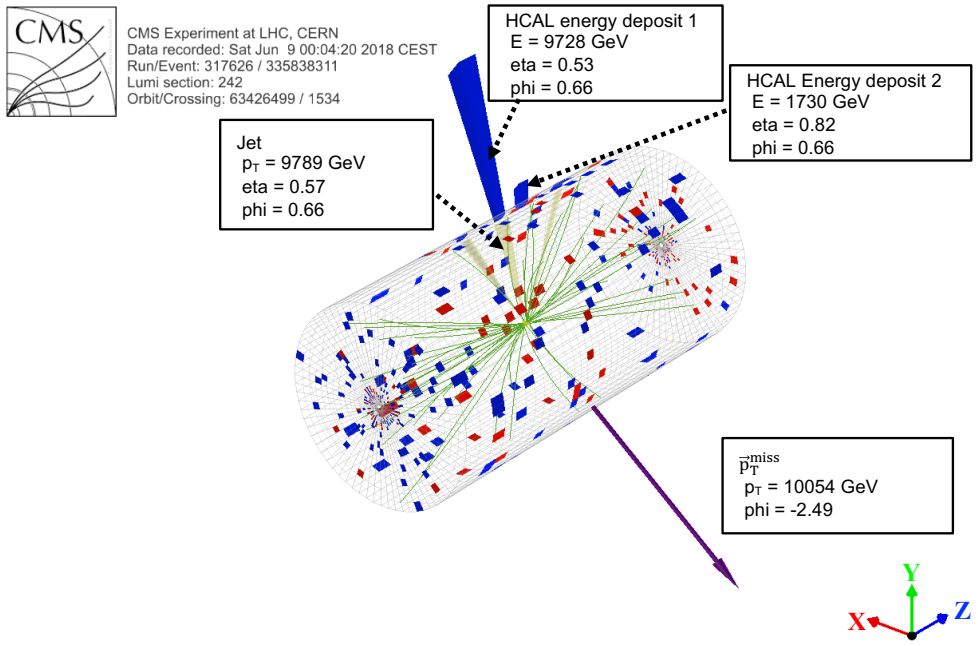


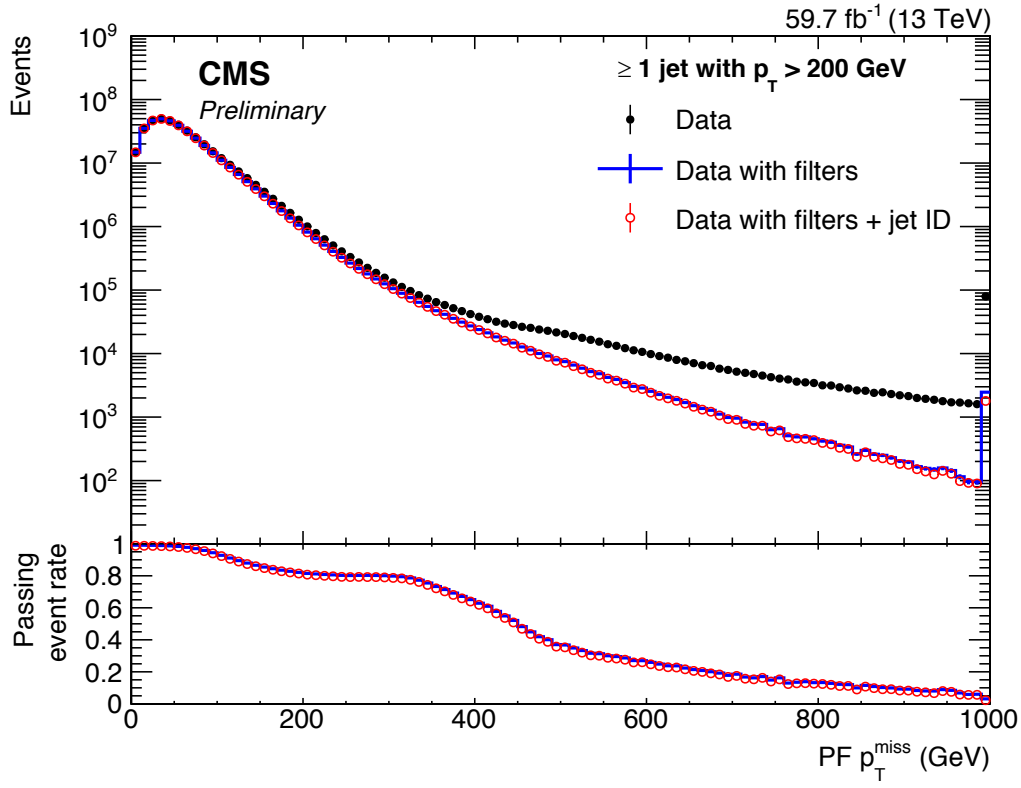
DP-2023/016 (TRG)

Event filters

- MET scanning team follows up on anomalous single events
- Appropriate MET filters are designed to keep out single-event failures:
  - cosmic muons, beam halo muons, calorimeter spikes, detector issues, reconstruction failures
- MET is very sensitive to any failure modes, so cleaning imperative

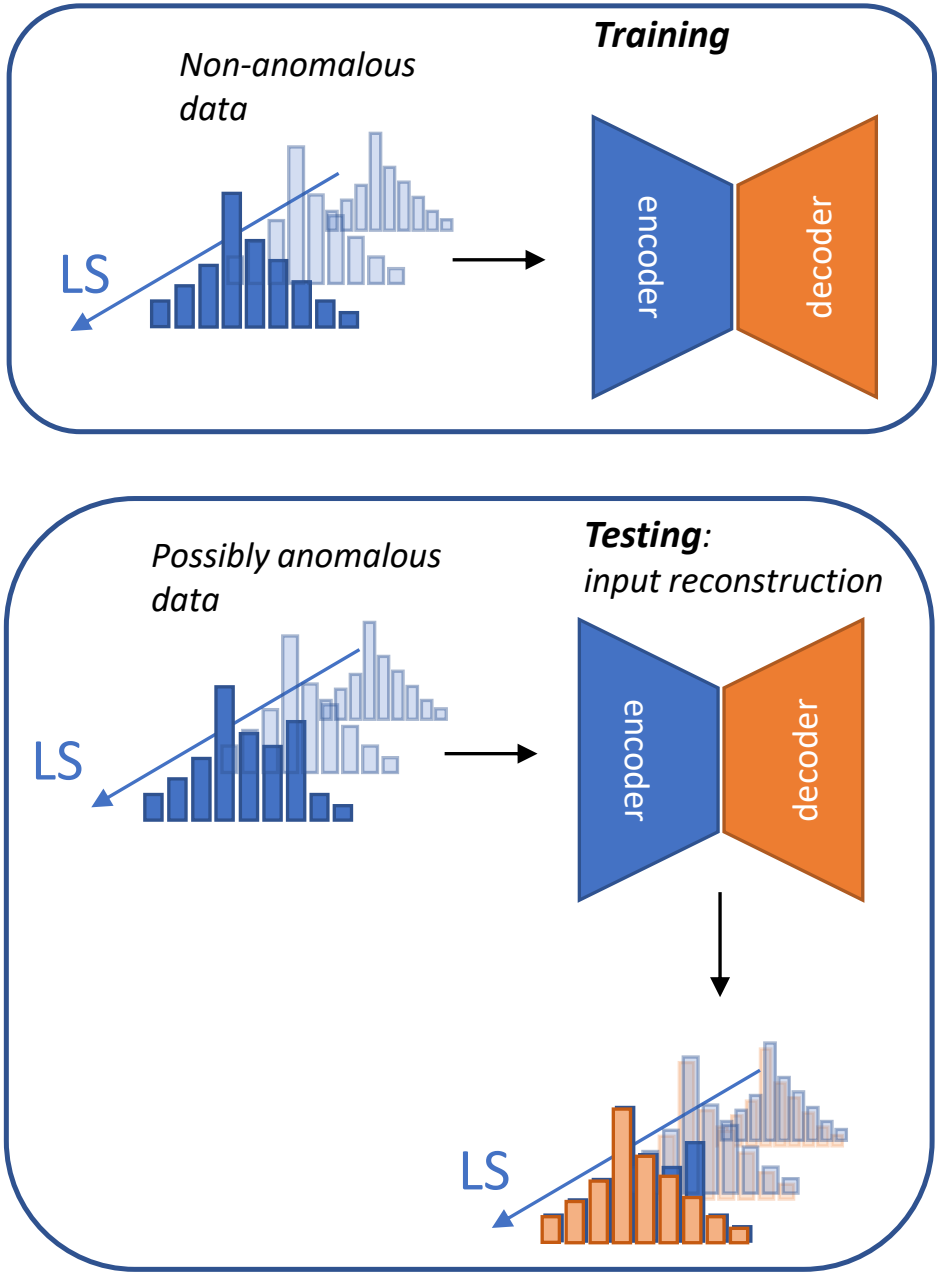
Noisy event due to HCAL detector related



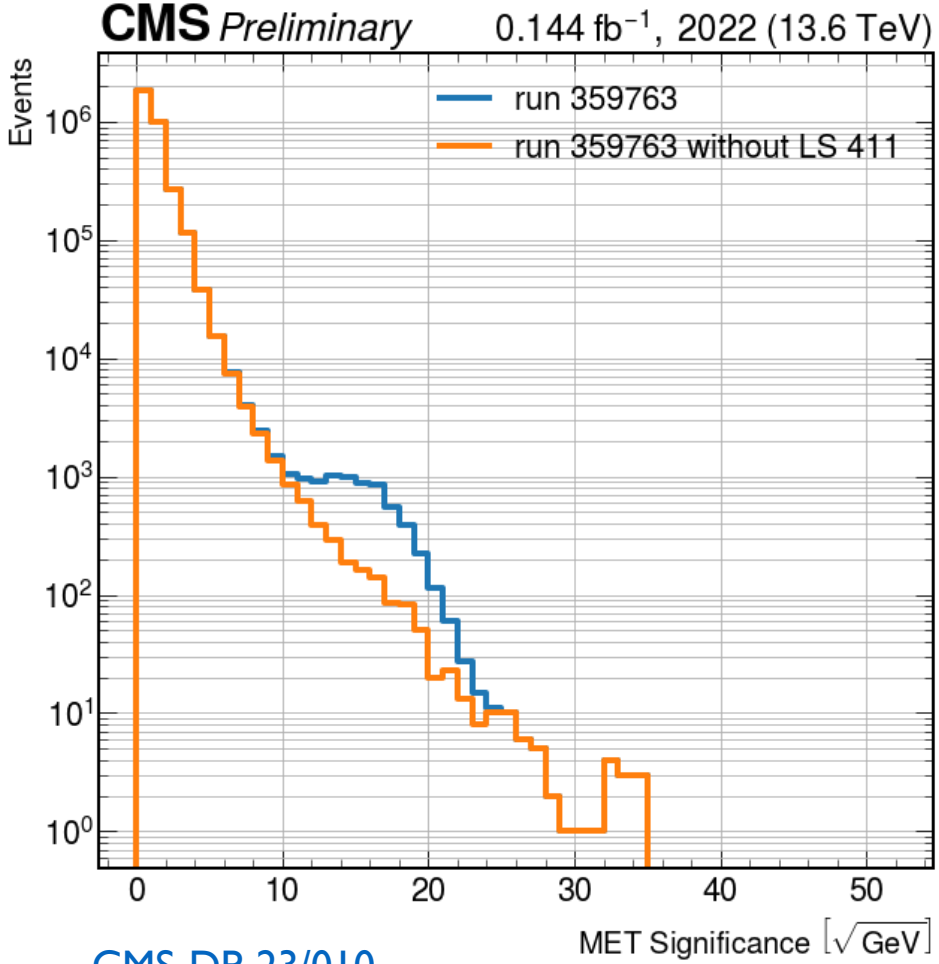


DP-2020/018

Data quality per section



- Data quality management (DQM) with Machine Learning (ML) improves overall efficiency:
  - AutoEncoder trained on good luminosity sections (LS ~ 23s) of data to learn normal behaviour
  - Anomalous data flagged per LS and removed if bad



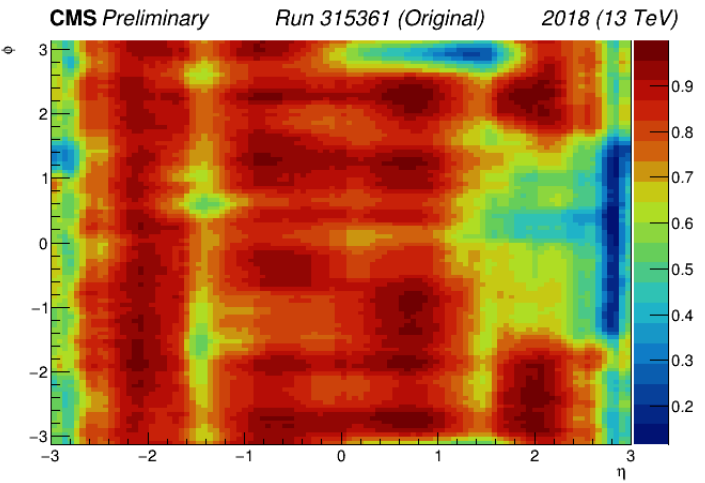
CMS-DP-23/010

Certification per run

- More subtle effects can be observed per run (typically hours of data instead of seconds)
  - Several machine learning techniques tested to assess run quality semi-automatically
  - Jet energy fractions most effective probes of data quality
- Only fully certified data must be used at higher levels (calibration, physics)

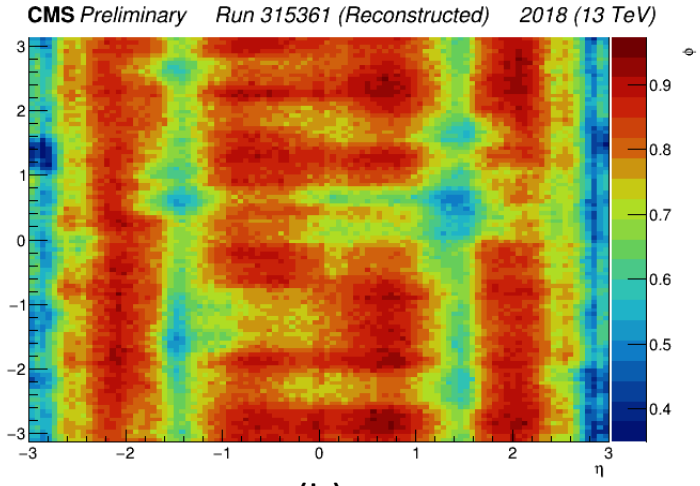
Hadron occupancy

Original



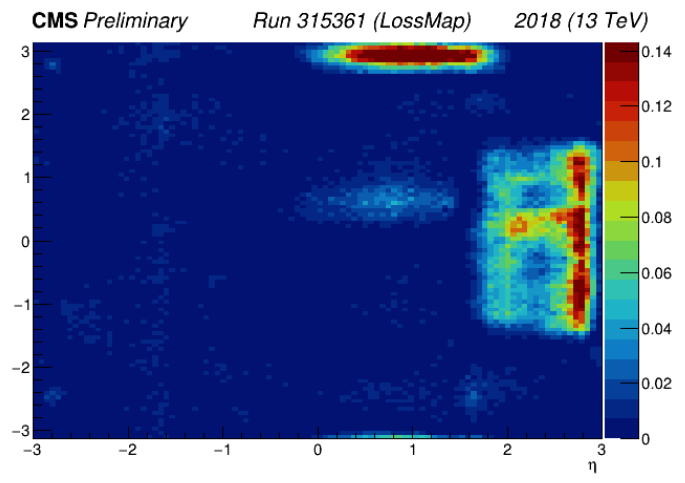
(a)

Reconstructed



(b)

Loss map

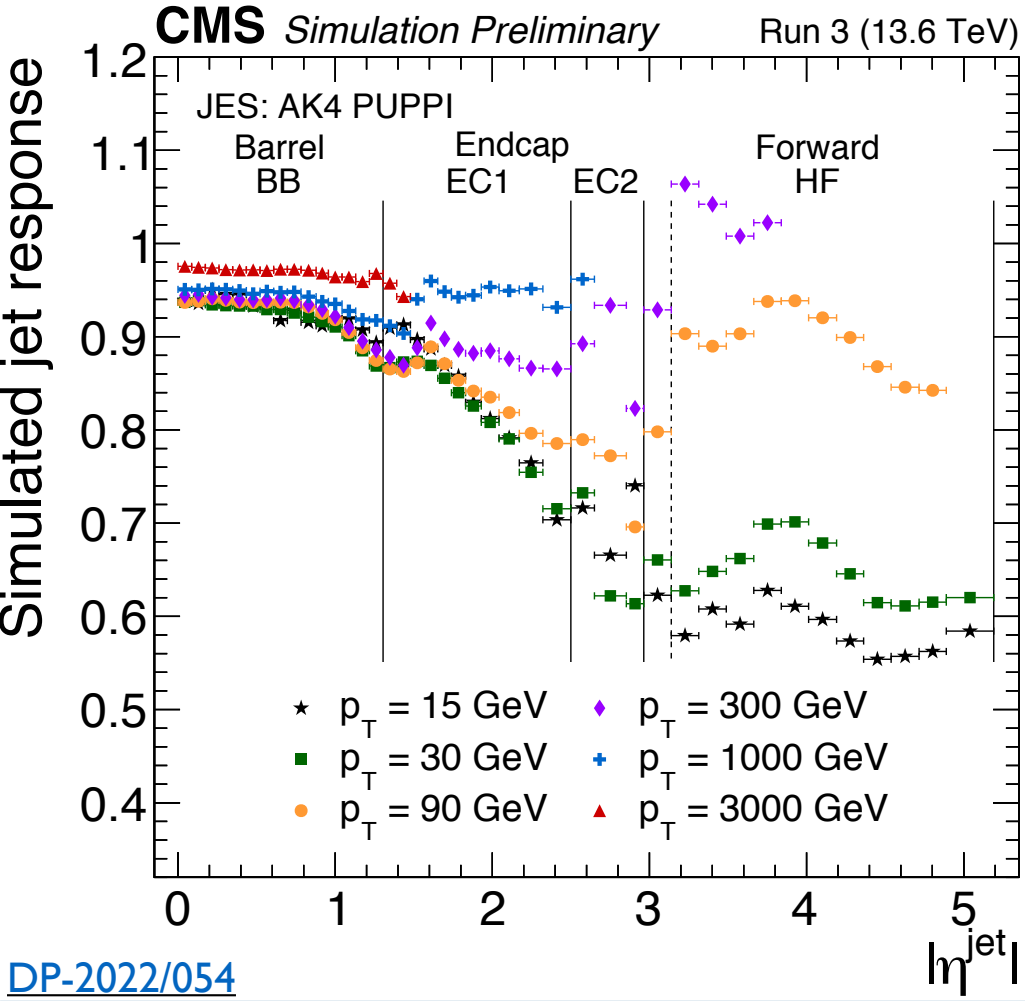


(c)

Simulated corrections

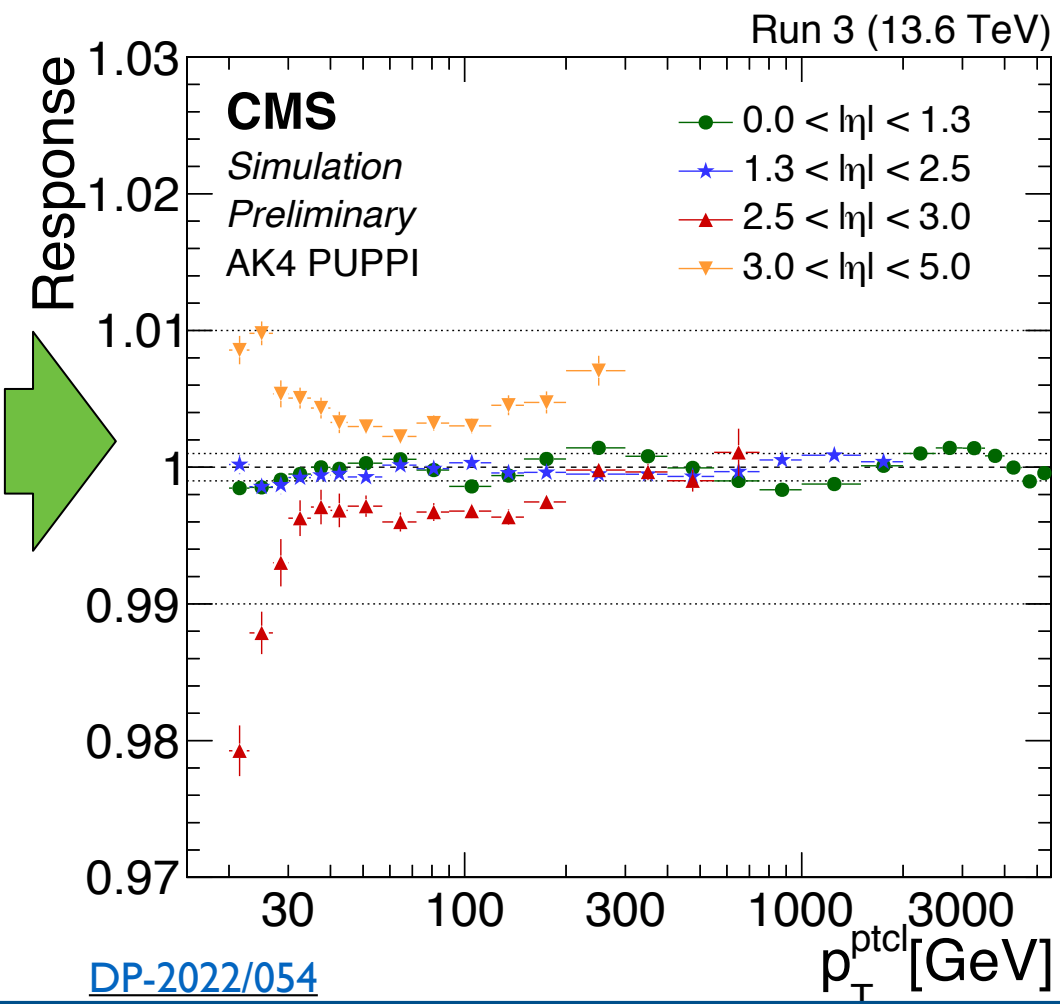
- Bulk of our calibrations are done with simulation, so it needs to be maintained very well
- Run 3: CHS -> PUPPI. Main impact for jet corrections: no separate pileup offset part
- Precision target: 0.1% within tracker coverage (|η|<2.5), ALARA elsewhere (<1%)

Simulated response



DP-2022/054

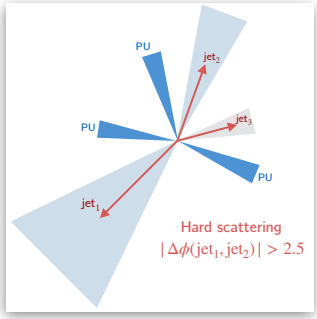
Simulation after corrections



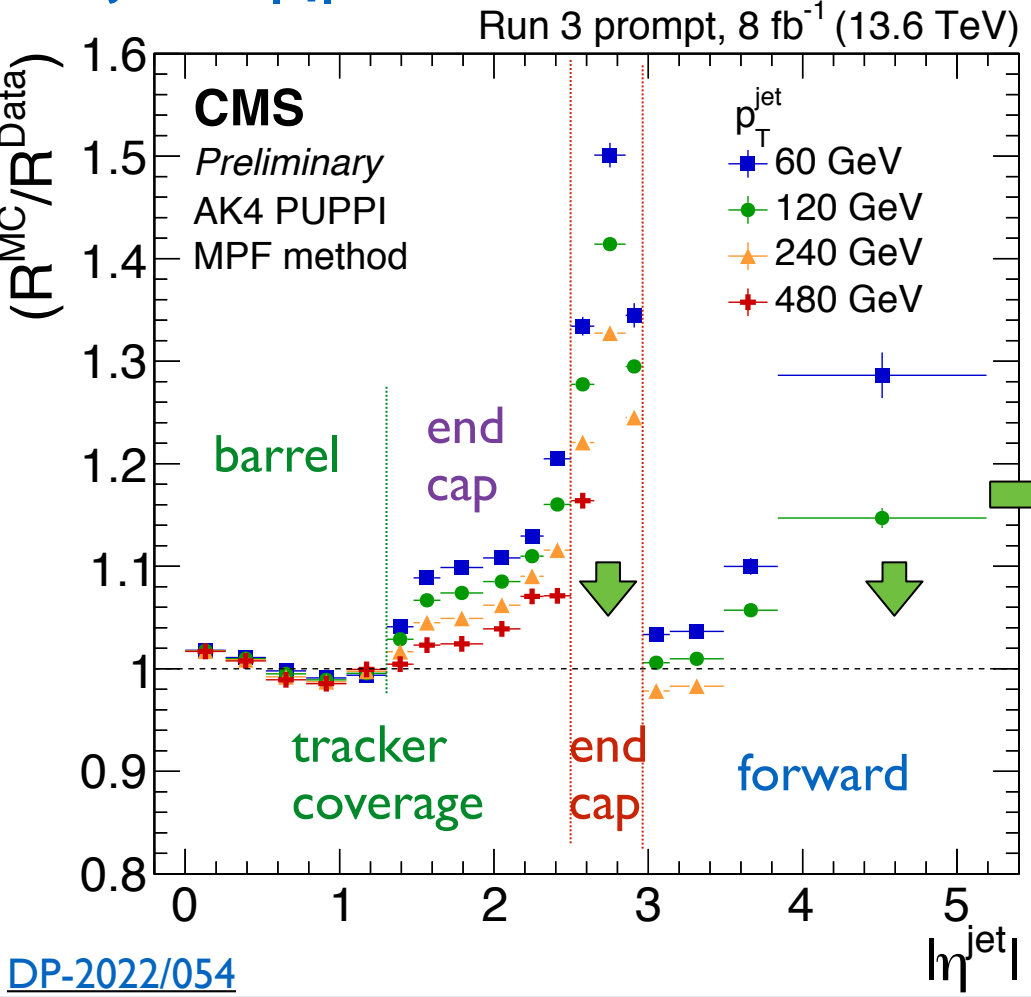
DP-2022/054

Data response

- Primary workhorse is dijet balance, to equalise response vs |η| (and p_T+time)
- Precision driven by global fit vs p_T: Z+jet, (W>qq'), γ+jet and multijet
- 0.1% goal: gluon radiation (FSR), unclustered energy, detector modelling

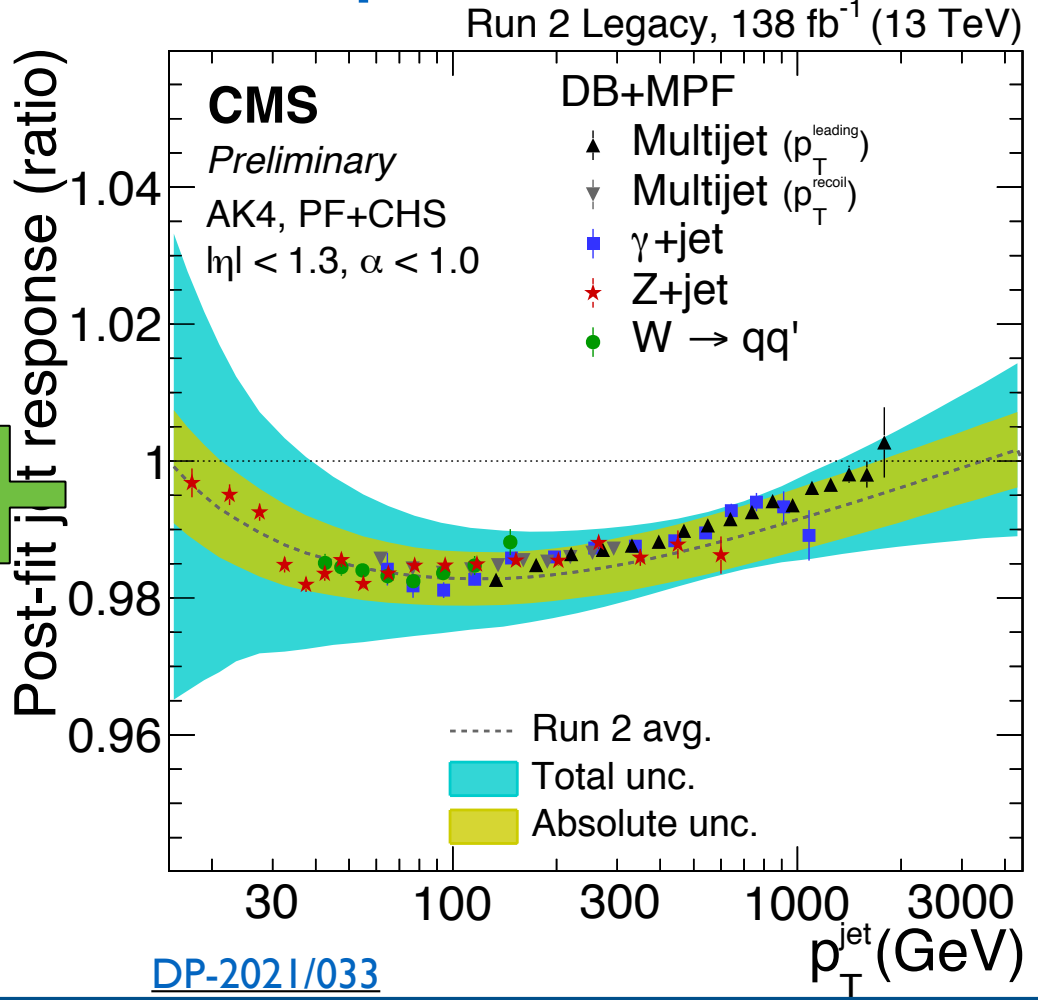


Dijet vs |η|



DP-2022/054

Global fit vs p_T

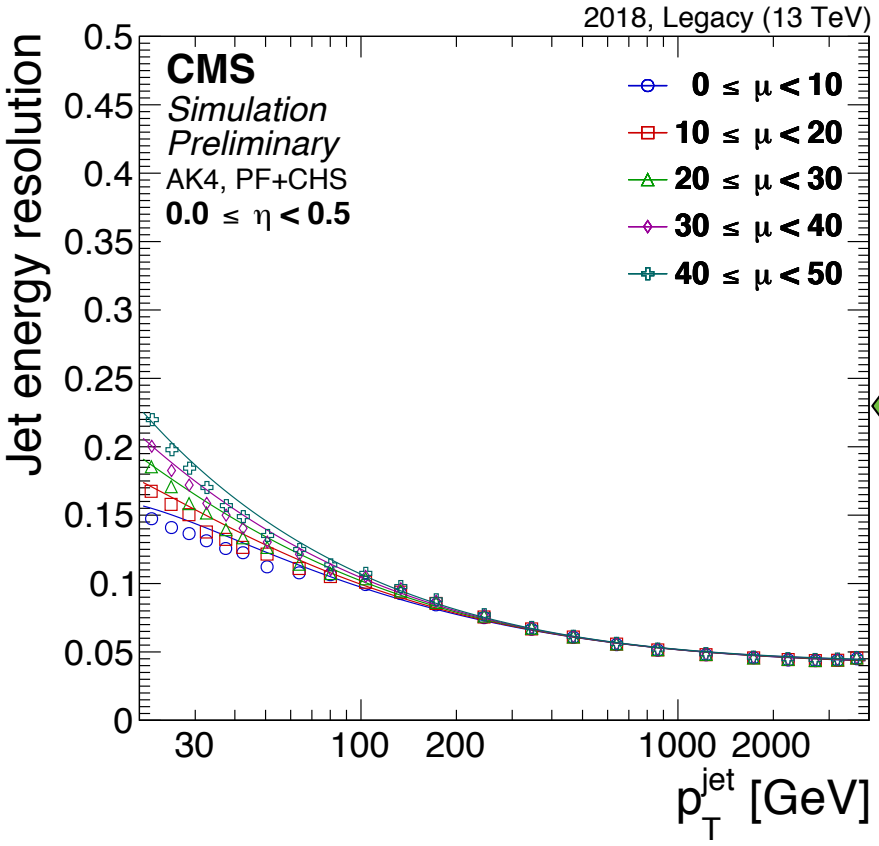


DP-2021/033

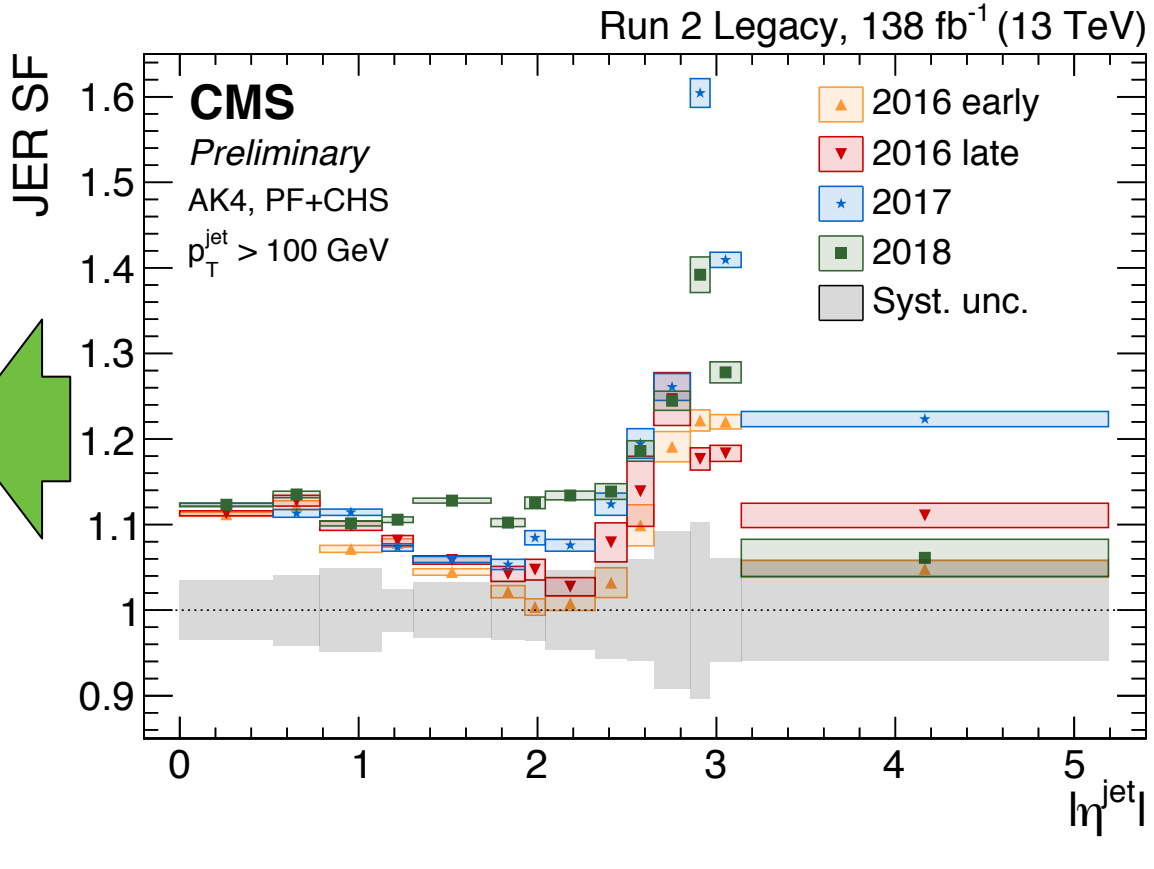
Jet p_T resolution

- For jet p_T resolution (JER) we mostly rely on simulation and pileup reweighing
- Simulation JER scaled to match data based on measurements from dijet balance
- In progress: fully p_T-dependent JER scale factors. Requires factorising detector effects

JER in simulation



Data/simulation for JER

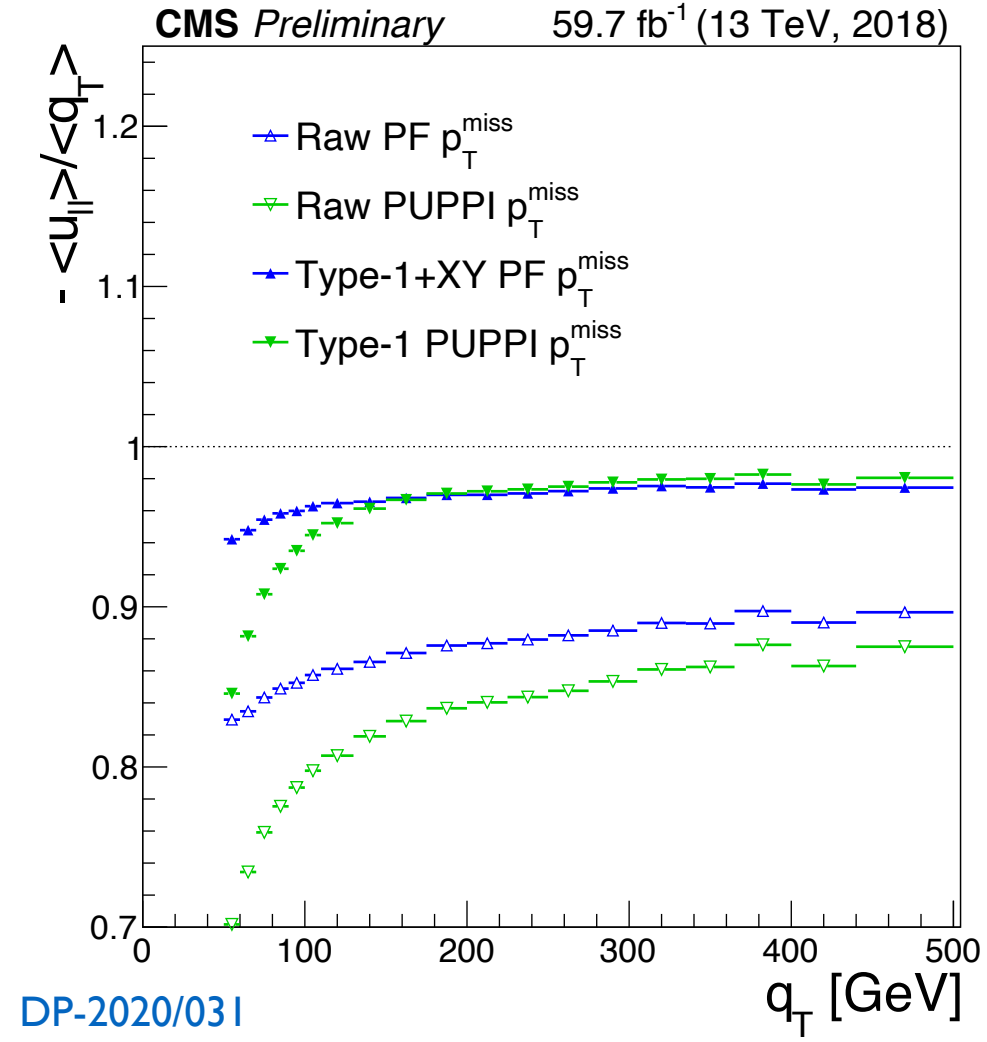


DP-2021/033

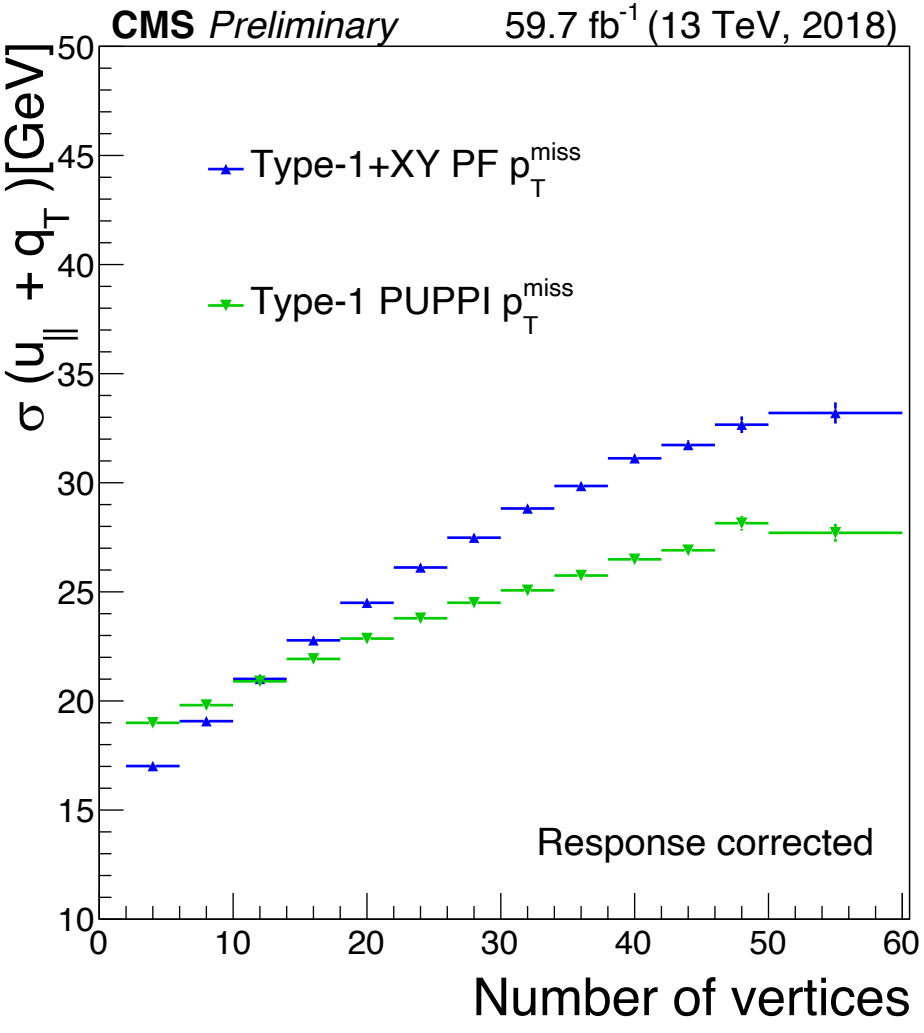
DP-2021/033

Missing momentum

- MET important to find neutrinos and massive weakly interacting particles
- Main metrics for comparisons are MET scale and MET resolution
- Future progress is in applications of ML to MET

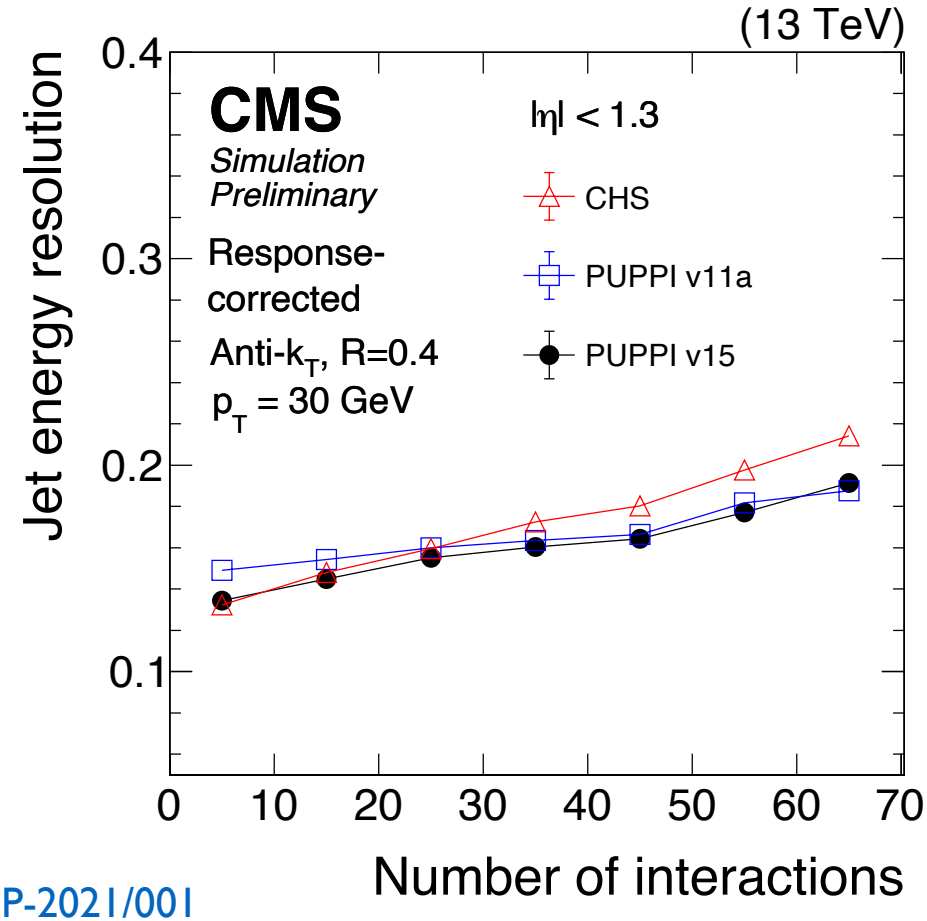


DP-2020/031

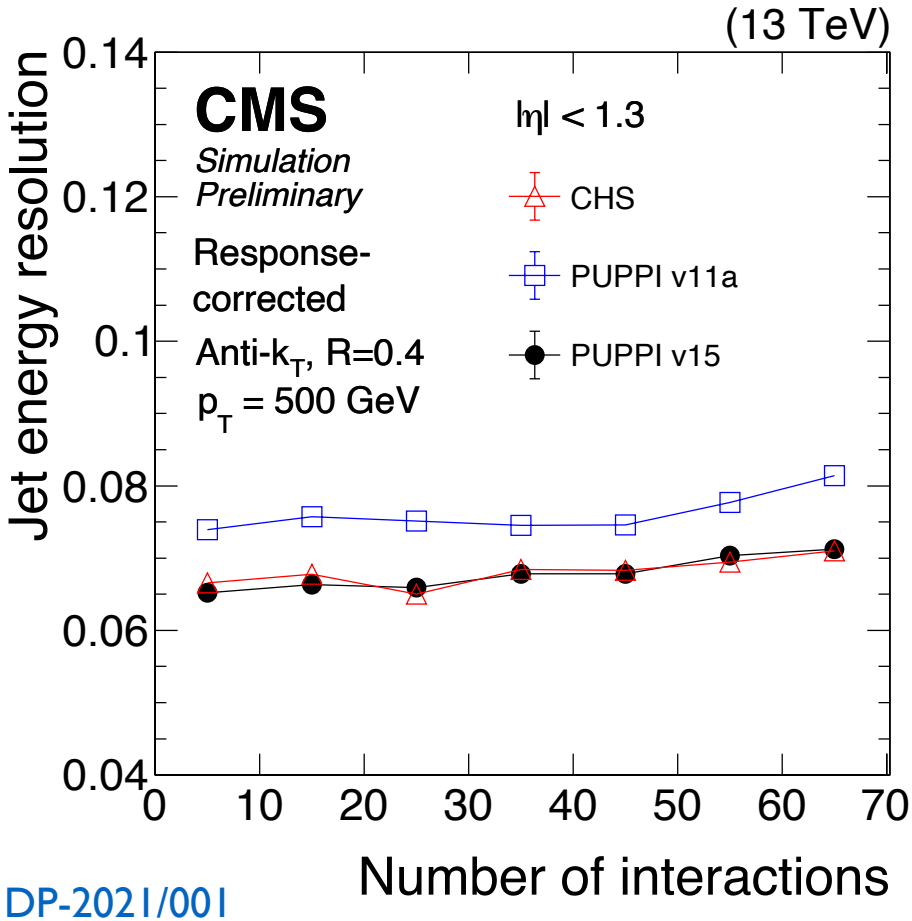


PUPPI tuning

- PUPPI is now at core of PF event reconstruction, but also requires tuning with feedback
- With appropriate choices, same or better than CHS essentially everywhere:
  - at low PU
  - at high p_T
  - in forward $|\eta|$



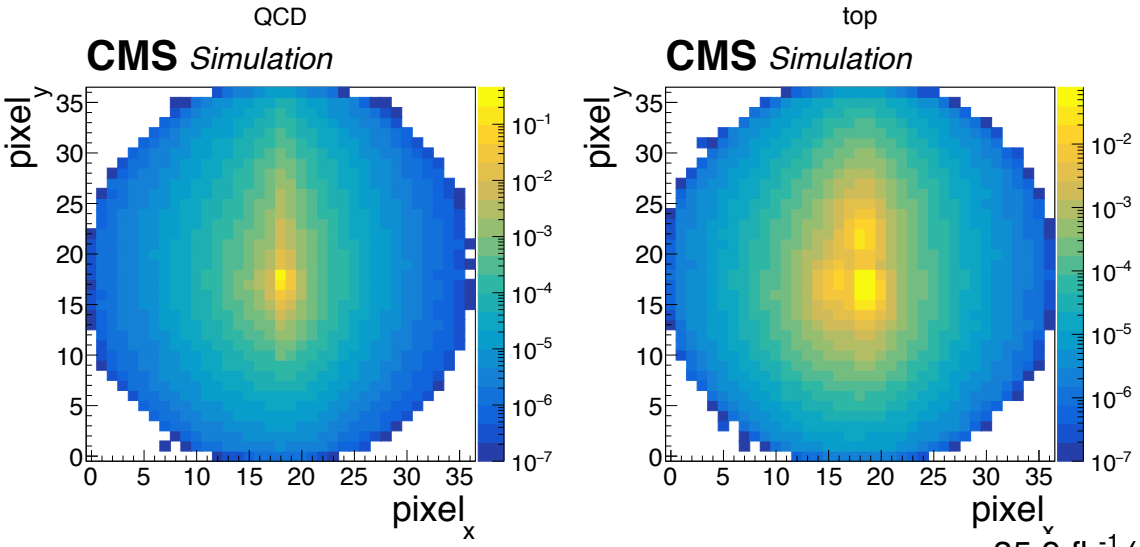
DP-2021/001

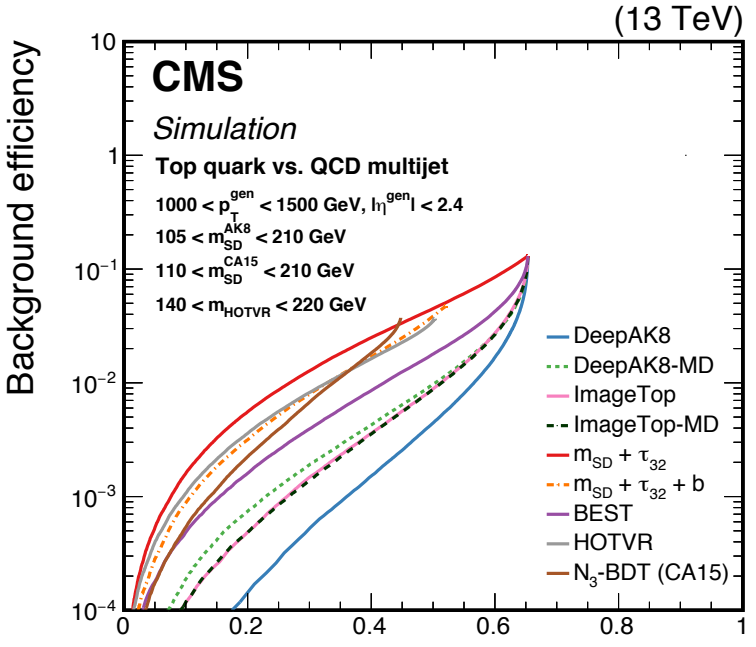


DP-2021/001

Top tagging

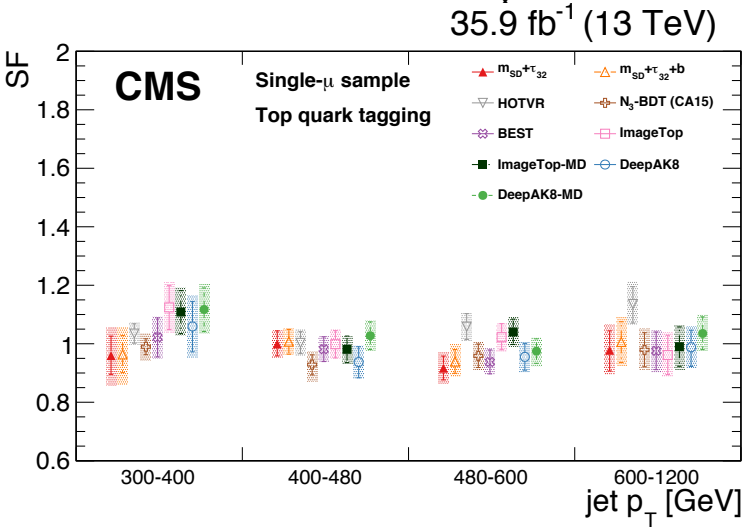
- Top (t>bqq') and other heavy (W,Z,H) resonances can be tagged with jet substructure
  - in simplest case, looking at image of jet to see 2–3 sub-jets
  - sub-jet mass another highly discriminating observable
  - eventually, ML for best performance

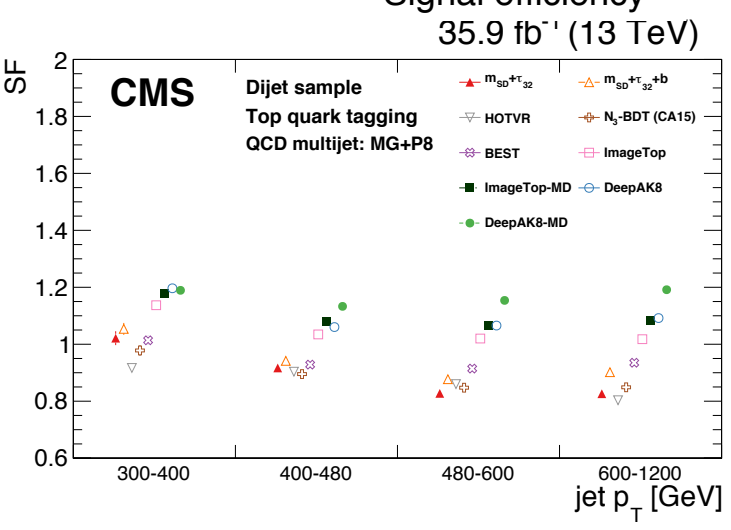




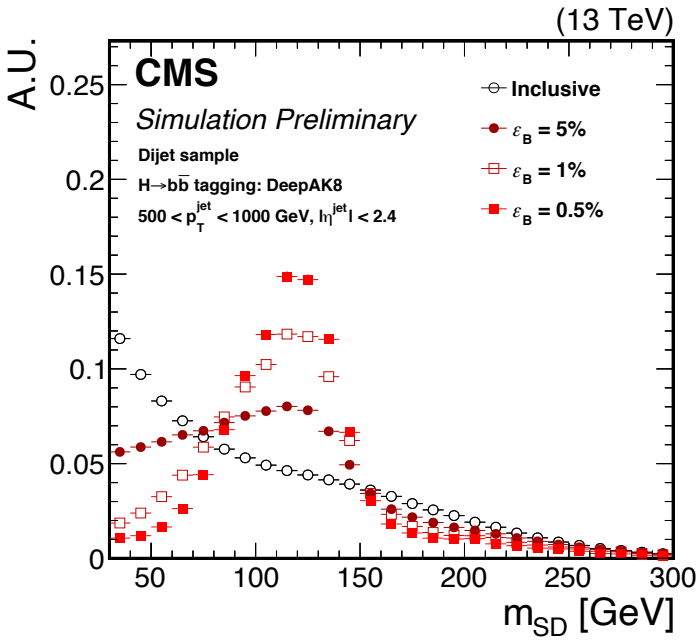
- Lot of work to characterise in data:
  - signal tagging efficiency
  - background mistake rate

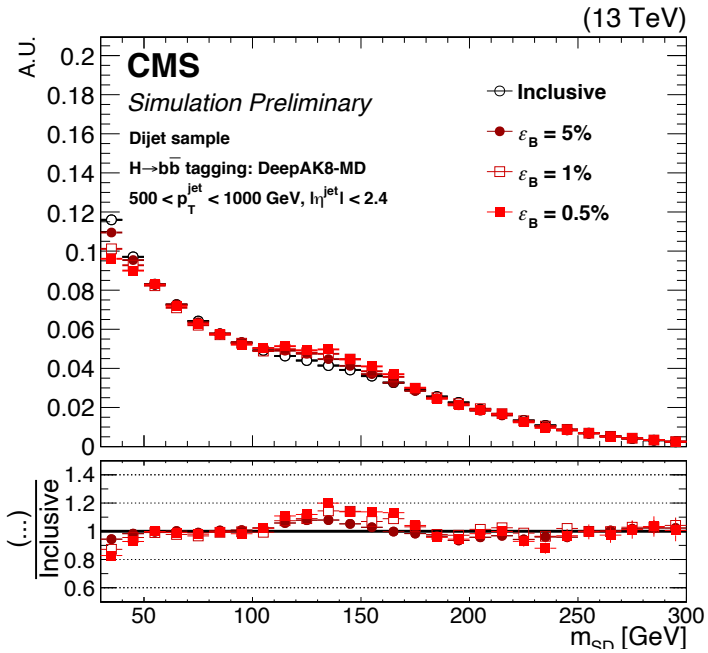
CMS-JME-18-002



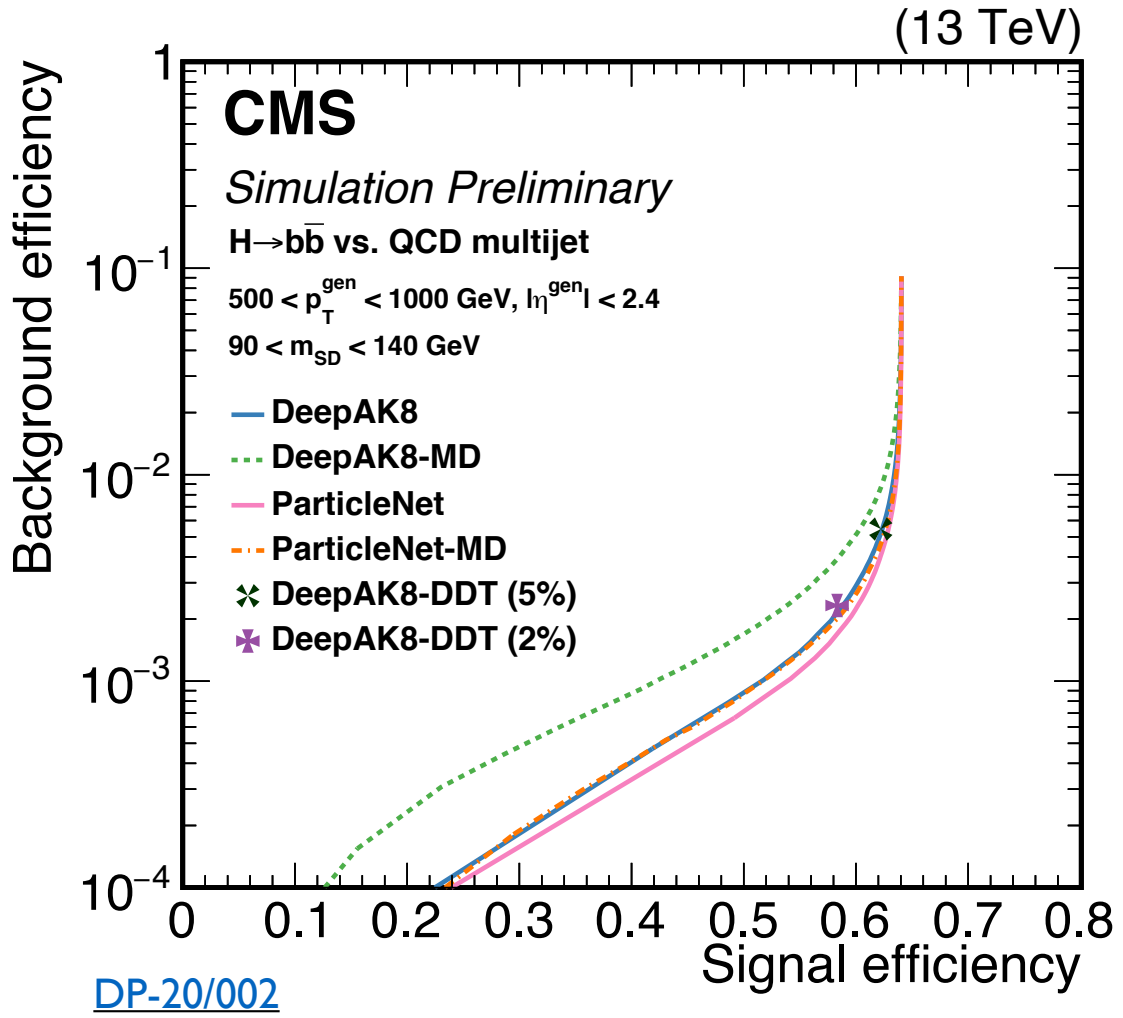


ParticleNet





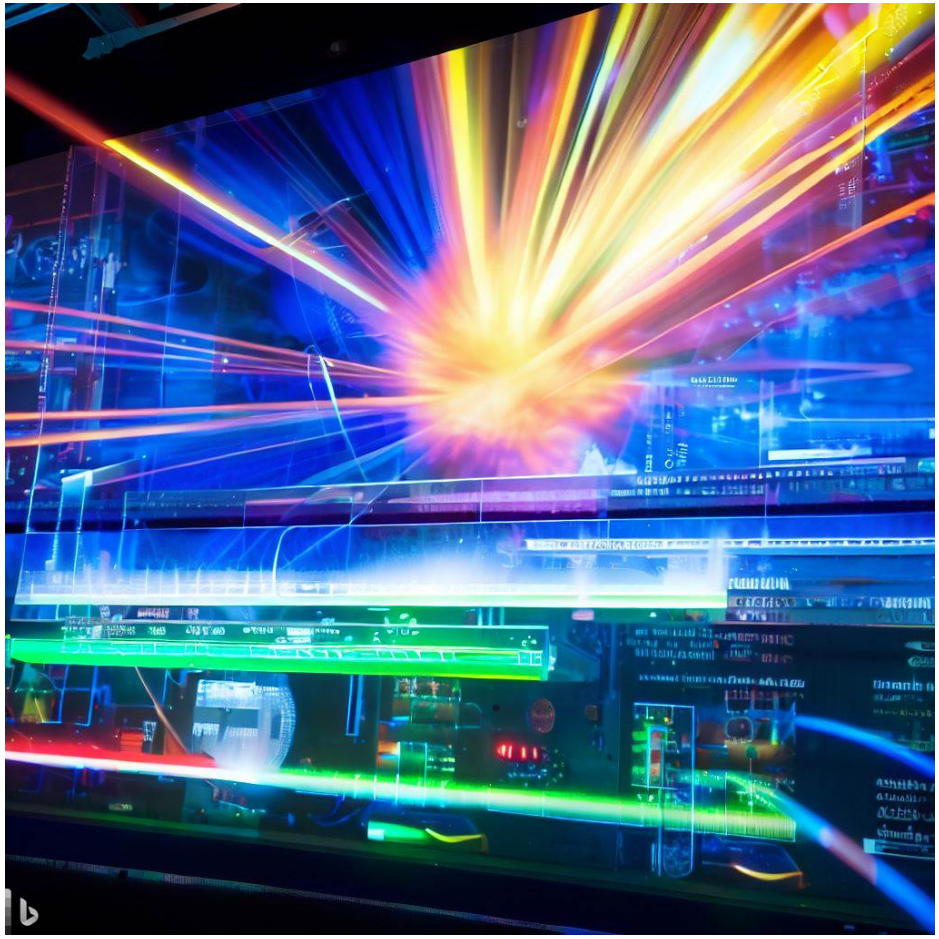
- Mass decorrelation key for taggers used in bump hunts
- ParticleNet based on graph neural network deep learning architectures now performance leader
- We will be seeing a lot more ParticleNet in the future



DP-20/002

Summary and outlook

- PUPPI now integral part of JetMET, full commissioning advancing well
- Machine learning key technique in future:
  - low-level reconstruction
  - high-level feature extraction
  - student recruitment



Backup