Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Emmy Noether-Programm
DFG Deutsche Forschungsgemeinschaft

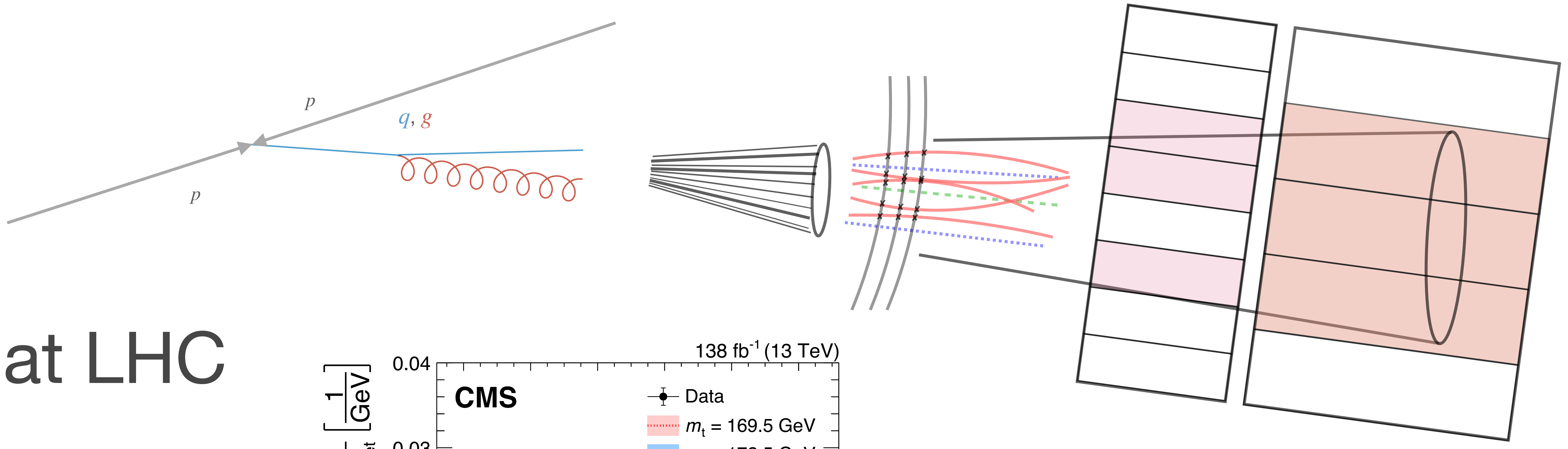
CMS

Jet performance and pileup mitigation in Run3 in CMS

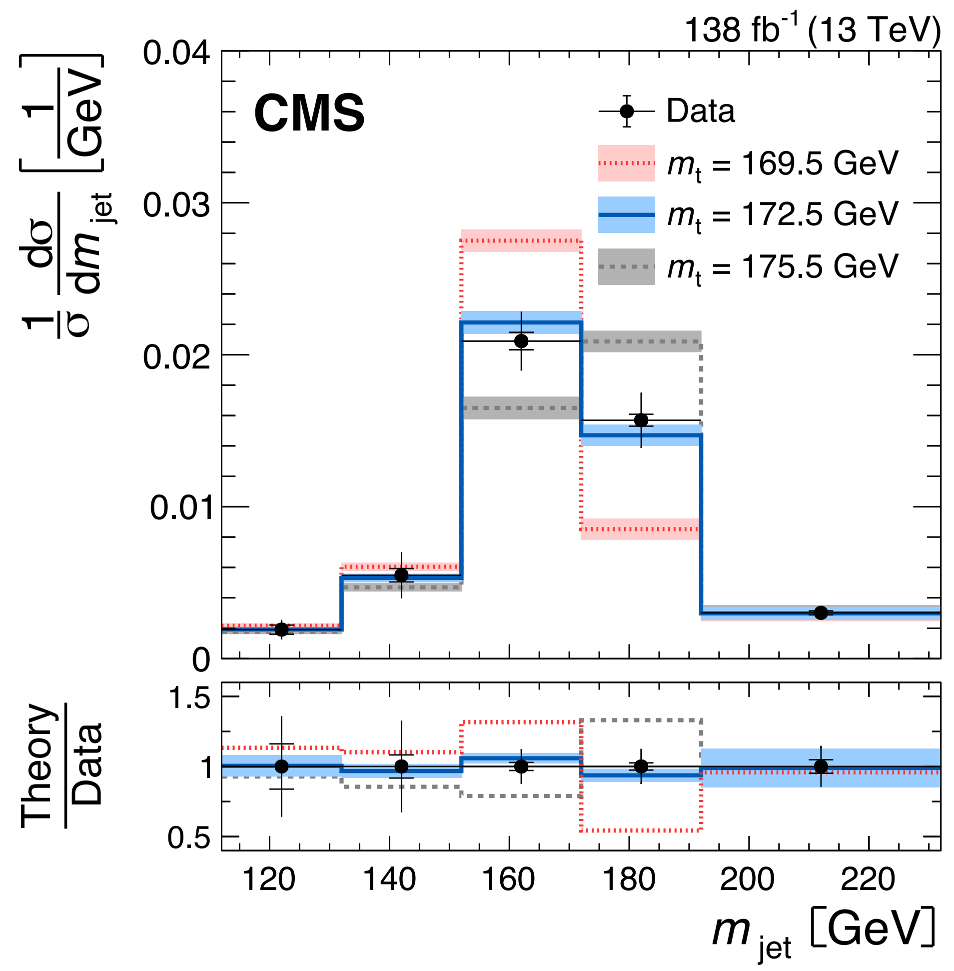
Steffen Albrecht
On behalf of the CMS Collaboration
July 31st 2023 - BOOST conference

Motivation

Jets in pp collisions

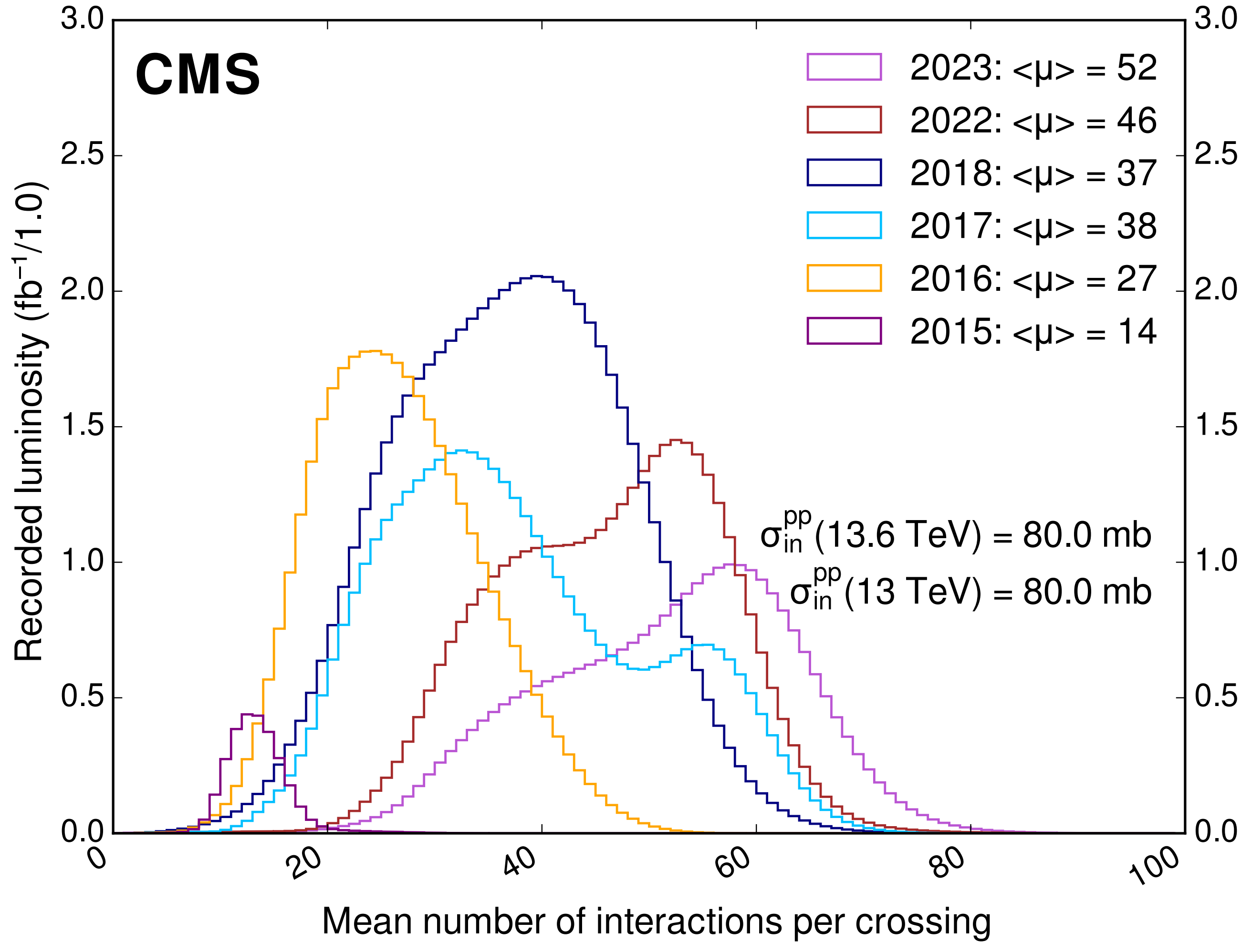


- Jets are produced in abundance at LHC
- Jets are crucial signatures in SM measurements and BSM searches
  - precision physics need precision calibration
  - Event pileup introduces challenges along the way
    - Run 3 so far: average 50 interactions per bunch crossing

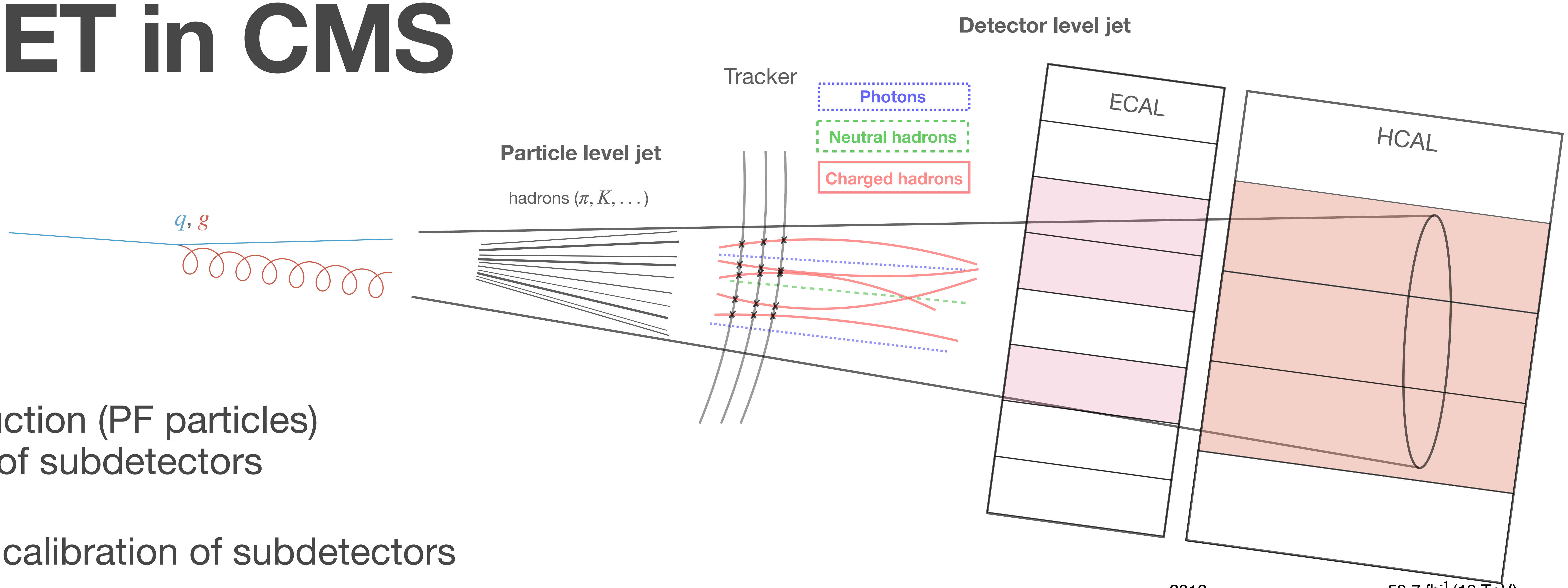


Eur. Phys. J. C 83, 560 (2023)

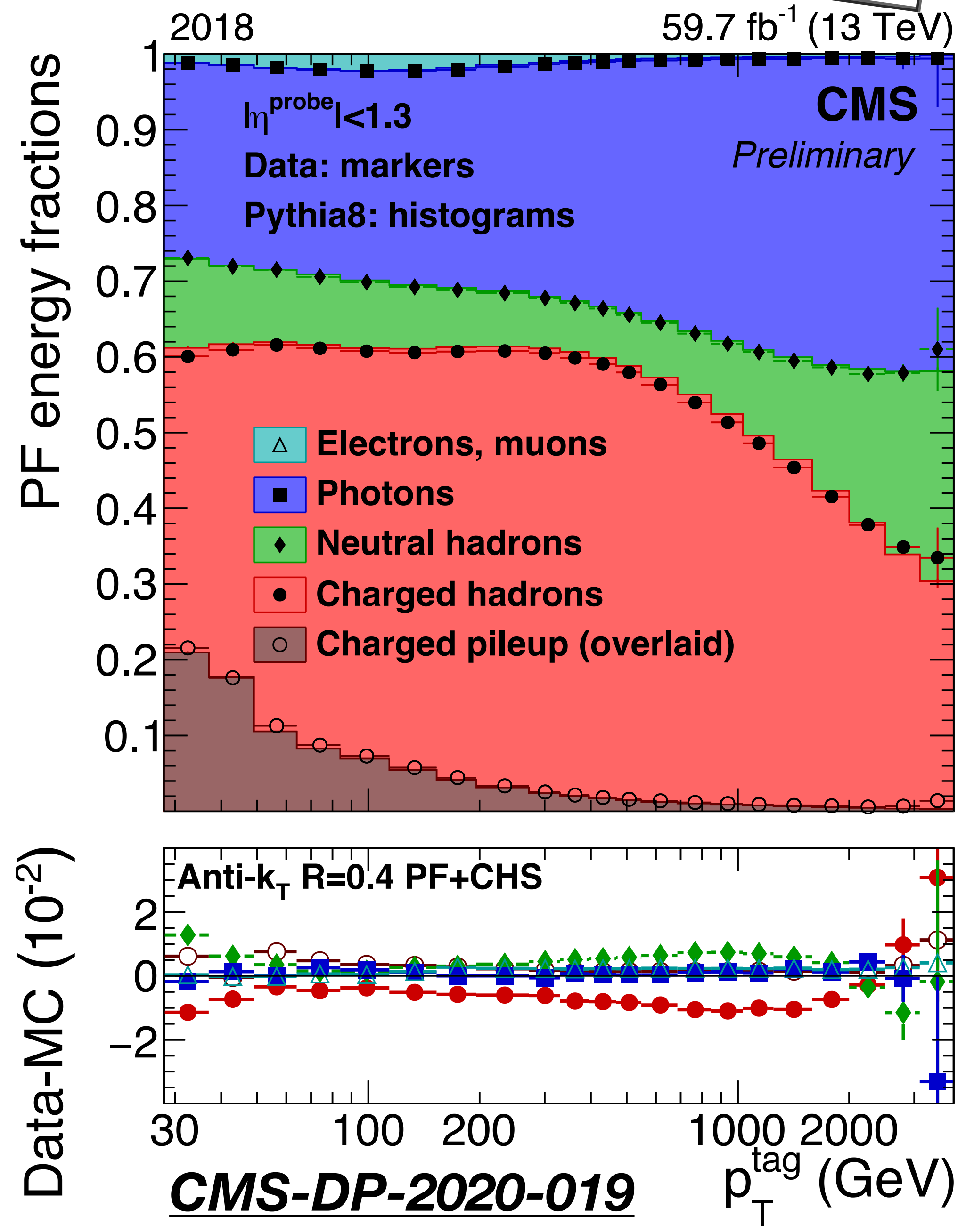
$$\Rightarrow m_t = 173.06 \pm 0.84\ \text{GeV}$$



Jets and MET in CMS

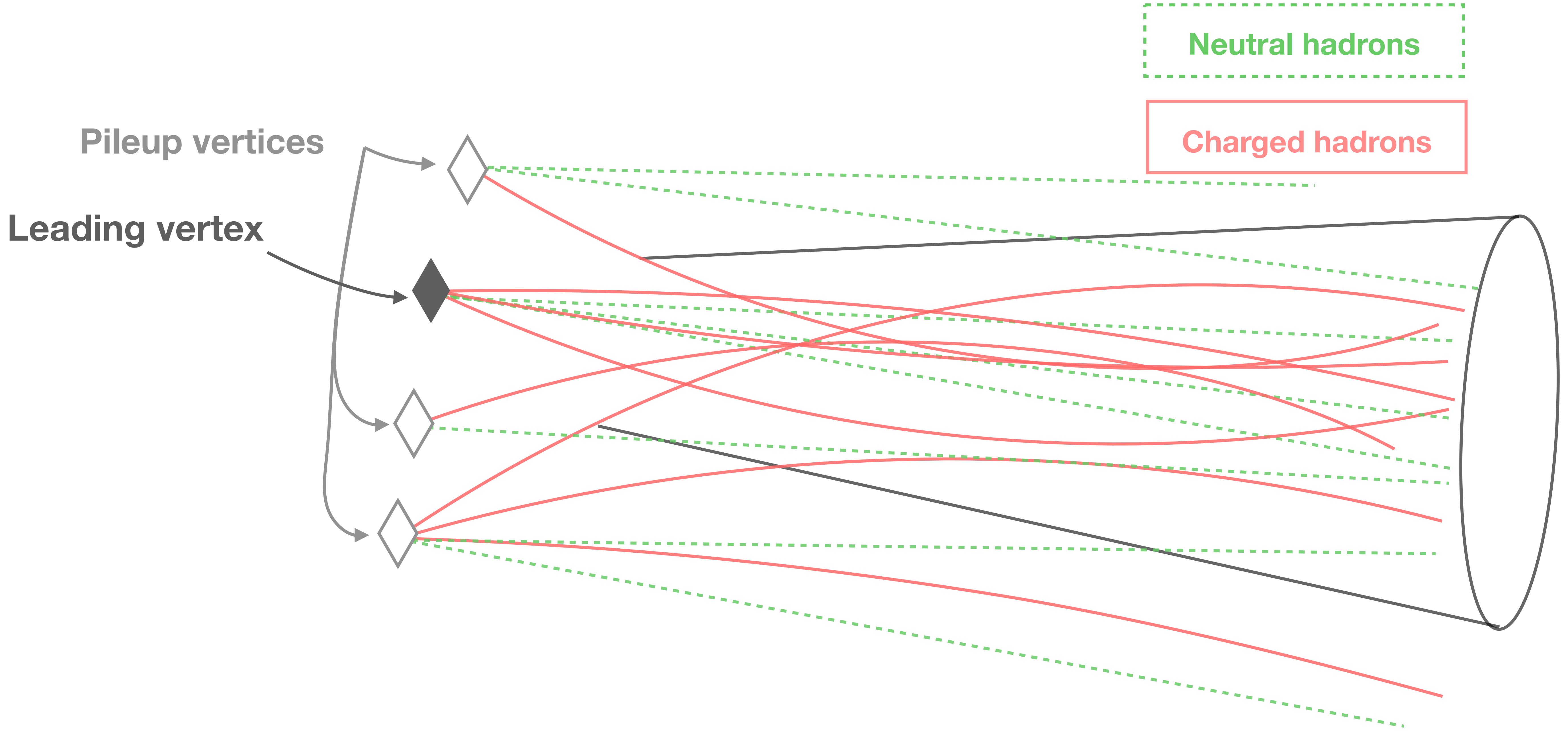


- ParticleFlow
  - local reconstruction (PF clusters,…)
  - global event reconstruction (PF particles) → linking information of subdetectors
    - Relying on precise calibration of subdetectors
    - Identification of particles (charged hadrons (tracker), neutral hadrons (HCAL), photons (ECAL),…)
- **Pileup mitigation (PUPPI)** →Jet clustering using anti-k_T (R=0.4, 0.8) → **Jet calibration in factorized approach**
- Also: MET, JetMET in HLT trigger and Data certification
- Not covered in this talk:
  - heavy-flavour tagging in boosted regime (→ see talk by Congqiao)
  - boosted jet tagging (→ see talk by Oz)



Pileup mitigation in Run 3

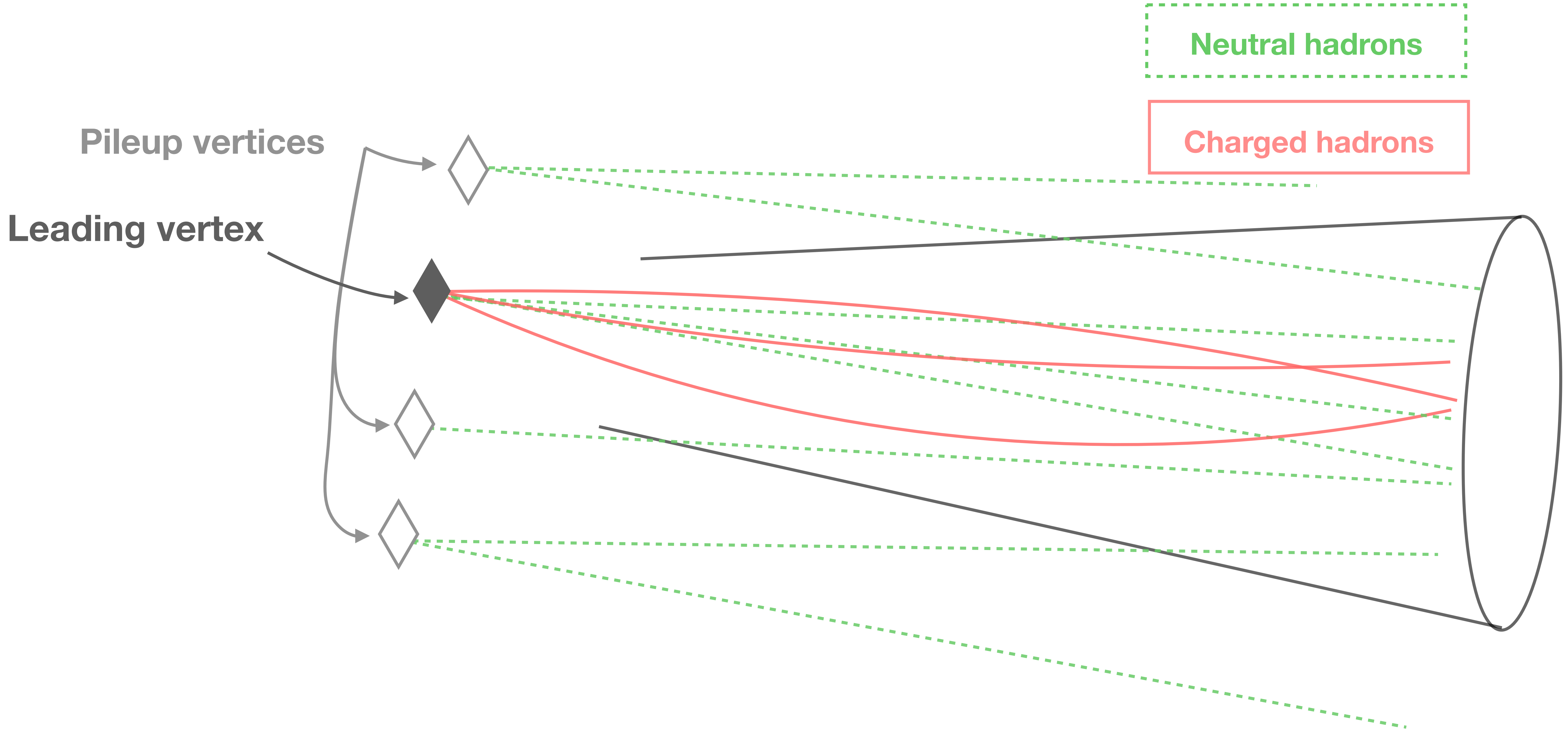
- Pileup interactions from same or other bunch crossing overlay the hard scattering
  - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far



- Default in Run 2: charged hadron subtraction **(CHS)** + average pileup offset correction
- Default in Run 3: treatment of charged and neutral hadrons using pileup per particle identification **(PUPPI)**

Pileup mitigation in Run 3

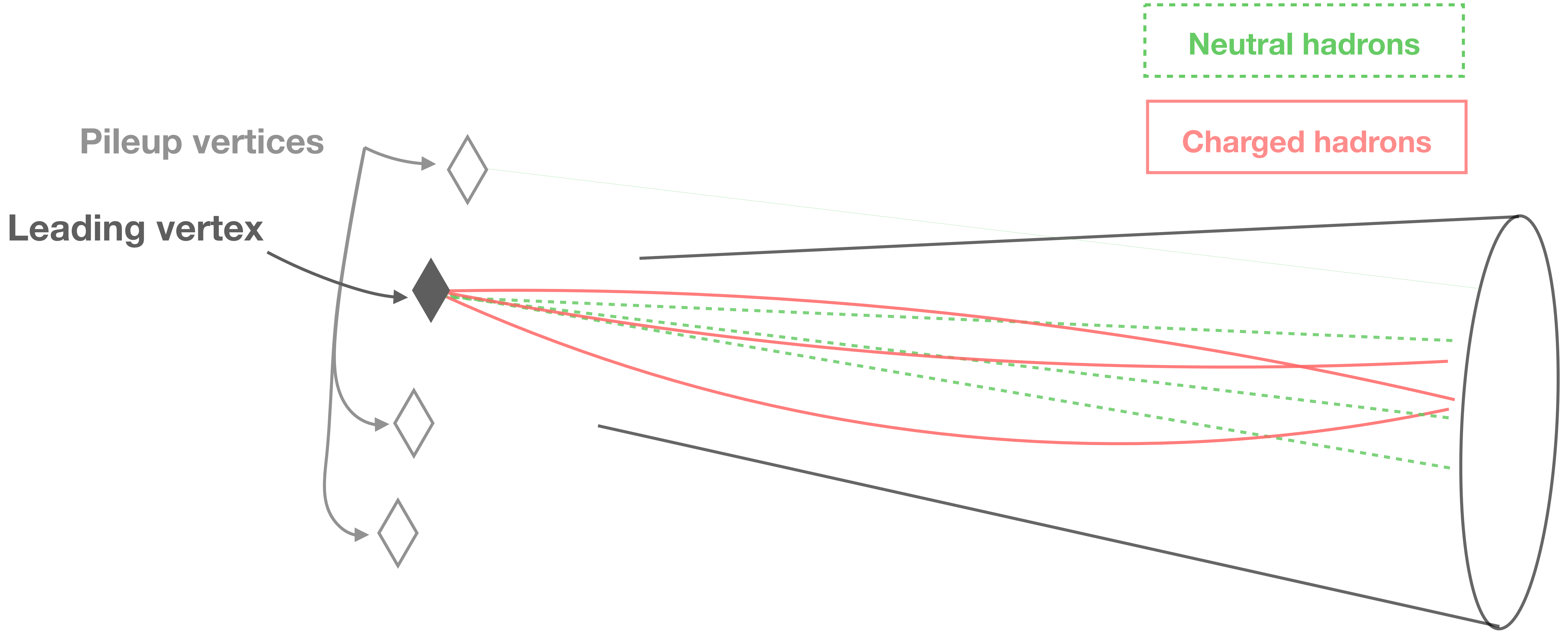
- Pileup interactions from same or other bunch crossing overlay the hard scattering
    - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far



- $\boxed{\textbf{Default in Run 2: } \text{charged hadron subtraction } \textbf{(CHS)} + \text{ average pileup offset correction}}$
- Default in Run 3: treatment of charged and neutral hadrons using pileup per particle identification **(PUPPI)**

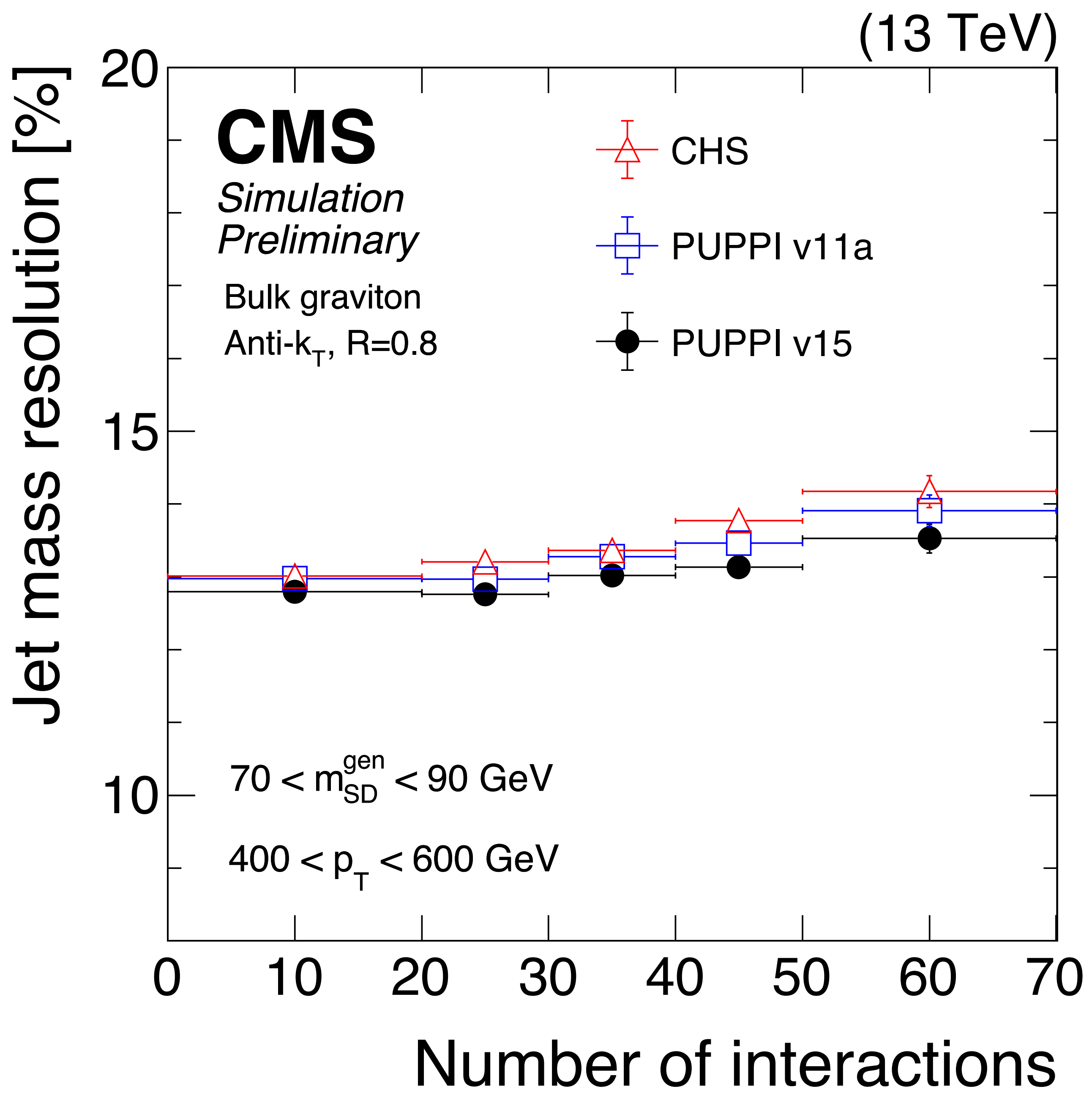
Pileup mitigation in Run 3

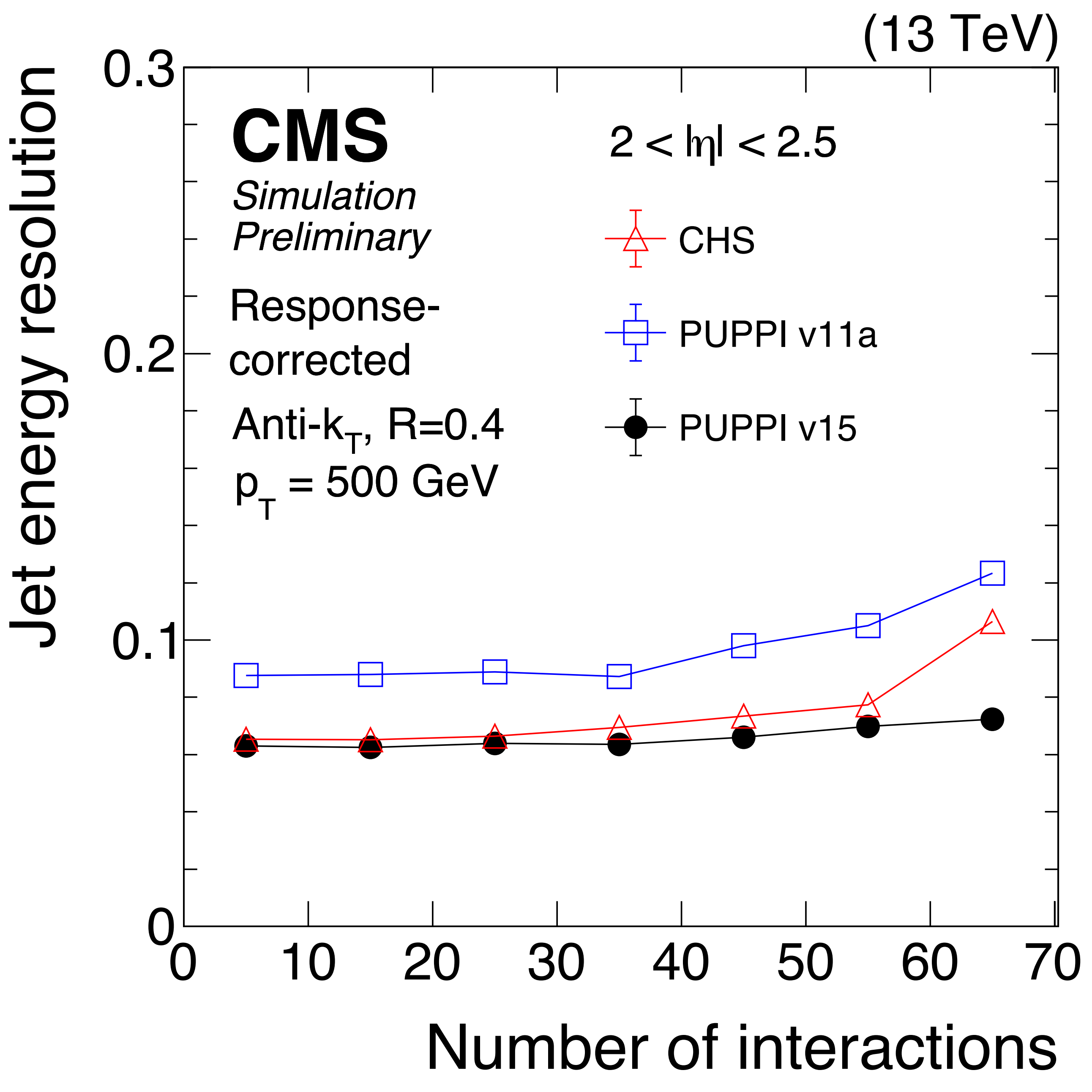
- Pileup interactions from same or other bunch crossing overlay the hard scattering
    - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far

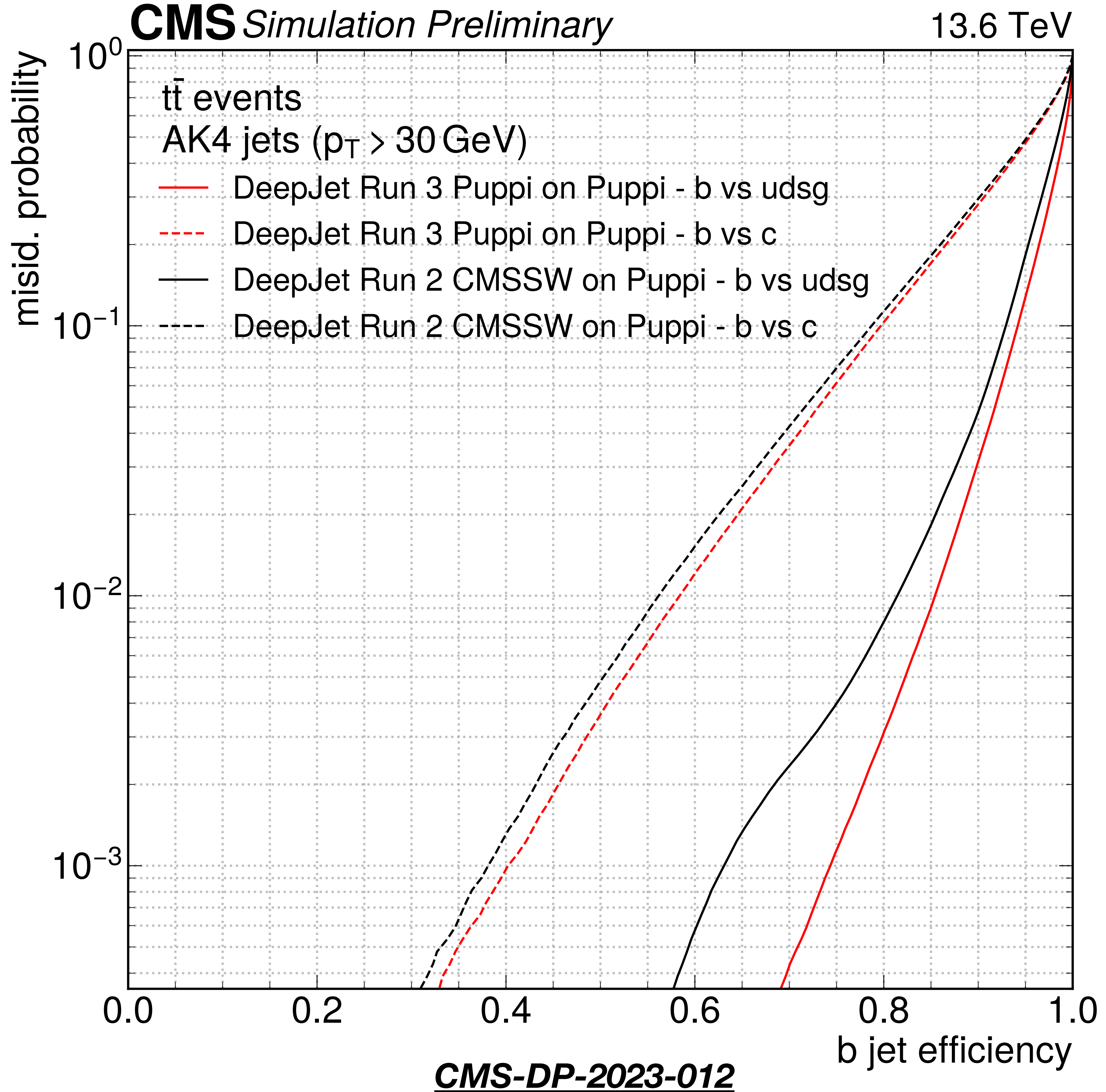


- Default in Run 2: charged hadron subtraction **(CHS)** + average pileup offset correction
- **Default in Run 3:** treatment of charged and neutral hadrons using pileup per particle identification **(PUPPI)**

Pileup mitigation in Run 3



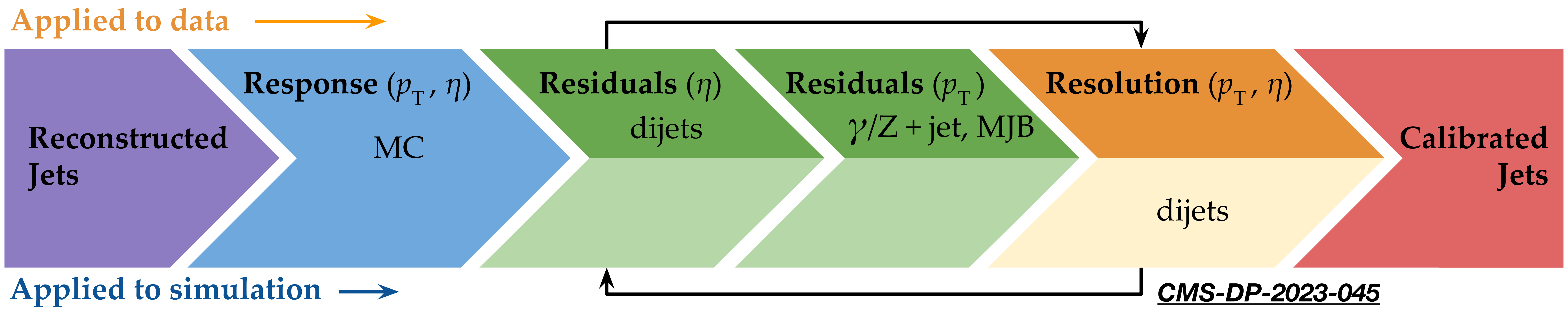




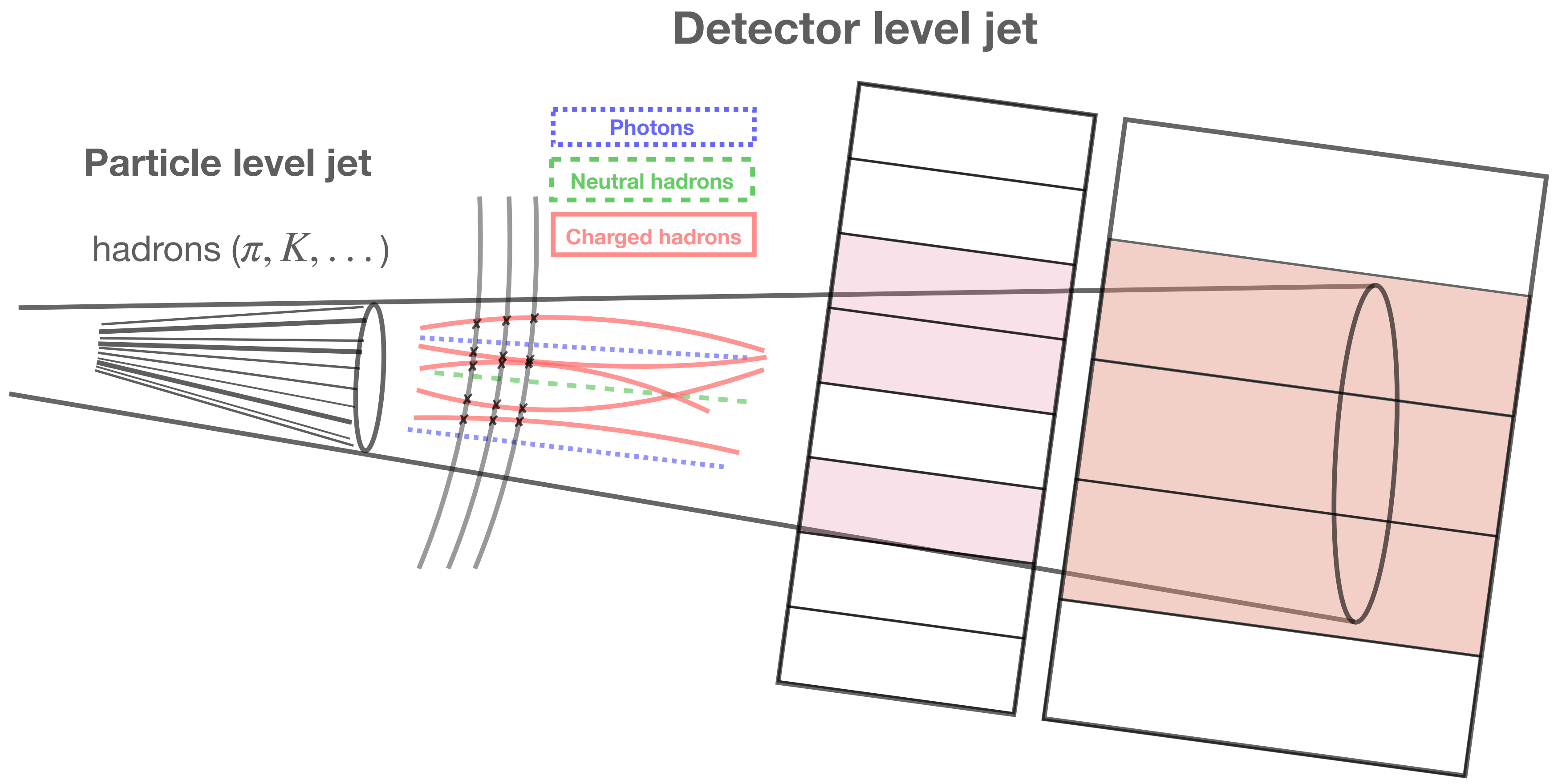
CMS-DP-2023-012

- PUPPI default method in Run 3 (using v15)
- PUPPI more stable against pileup in substructure variables
- New DeepJet training on AK4 PUPPI jets using Run3 data
    - Improved performance w.r.t. Run2 training (CHS in training)

Jet calibration in CMS



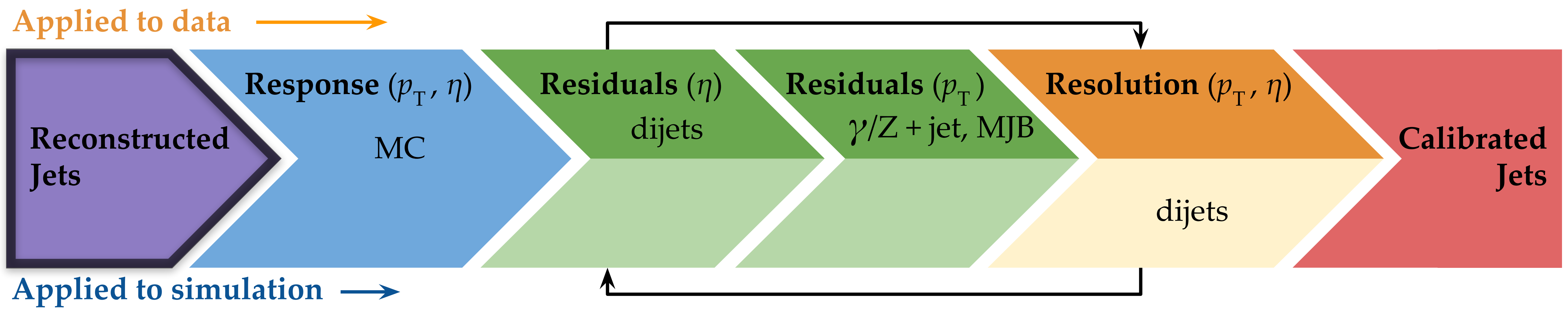
- Jet calibration in factorized approach to correct jet energy scale **JES** (in data and simulation) and jet energy resolution **JER** (in simulation)
    - MC truth corrections: correct JES to level of particle level jets
    - Residual corrections: correct JES for residual differences between data and simulation
    - Smear jet energy resolution in simulation to match that in data

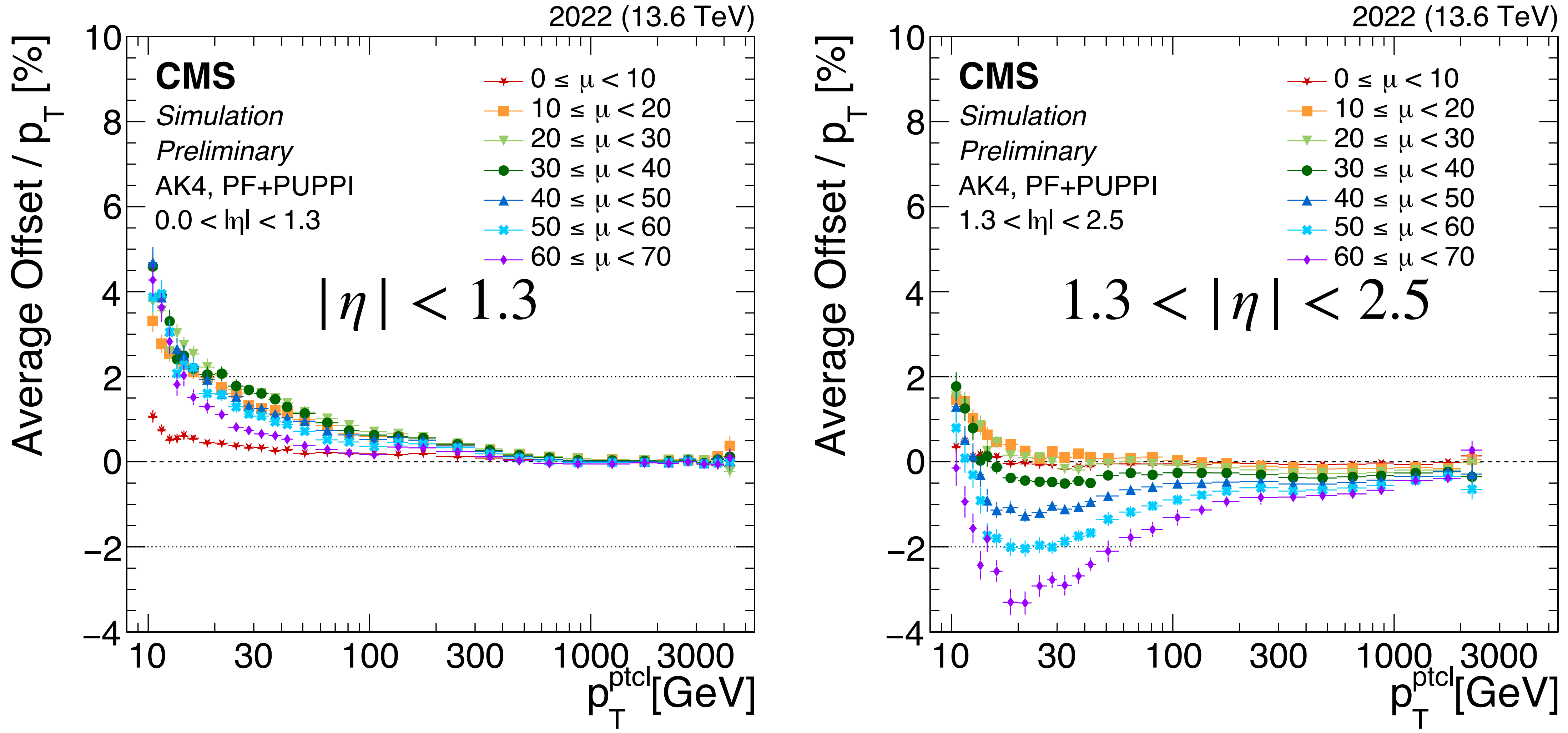


Jet calibration in CMS

Remnant pileup in jets

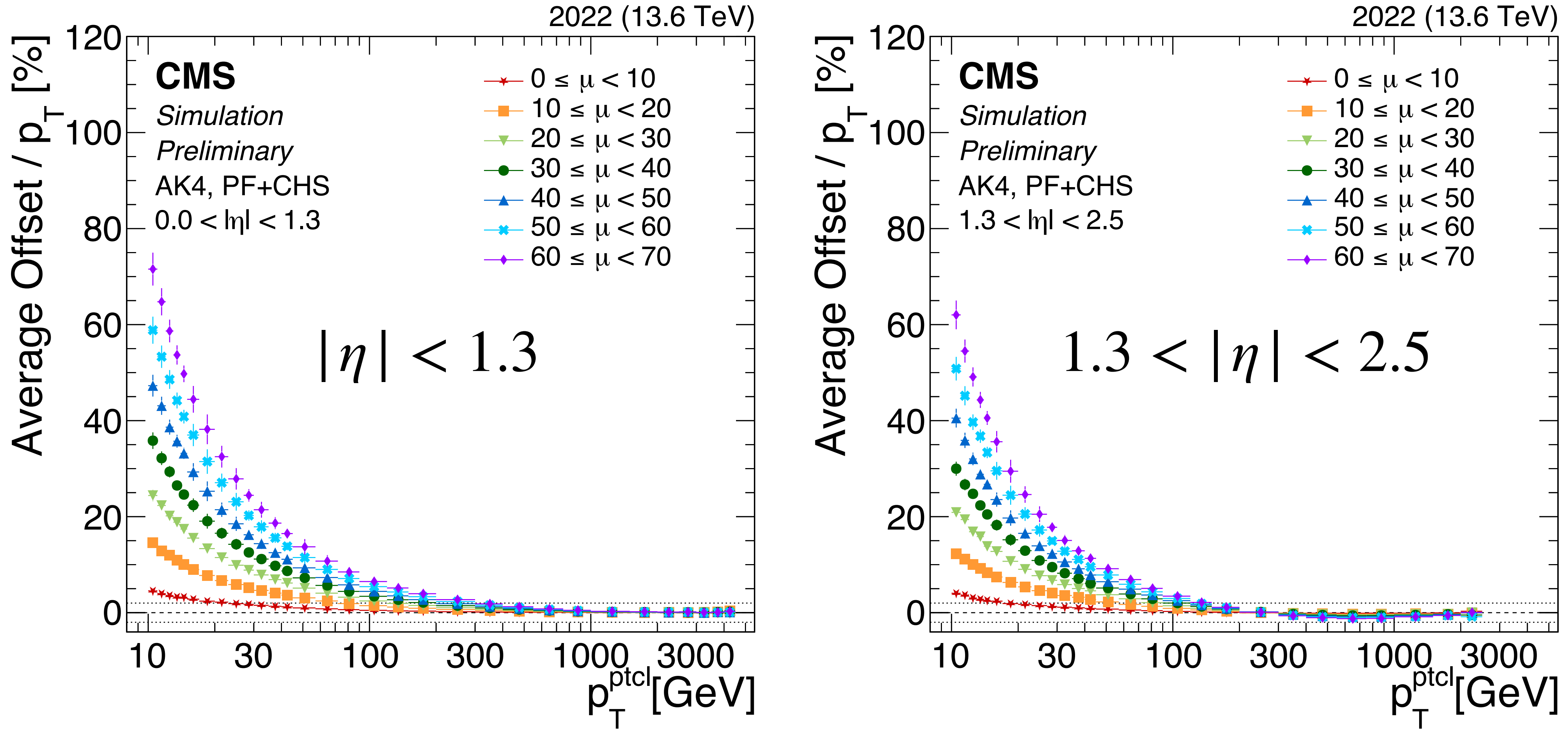
- Average pileup offset / $p_{T,\mathrm{ptcl}}$
- PUPPI jets enter JES/JER calibration workflow with minimal dependence on <μ>
  - For |η|<2.5 (barrel and endcap):
    - CHS jets have (without offset corrections) up to ~70% remnant pileup
    - PUPPI jets <4% for p_T>20 GeV
- No further pileup corrections applied to PUPPI jets





PUPPI

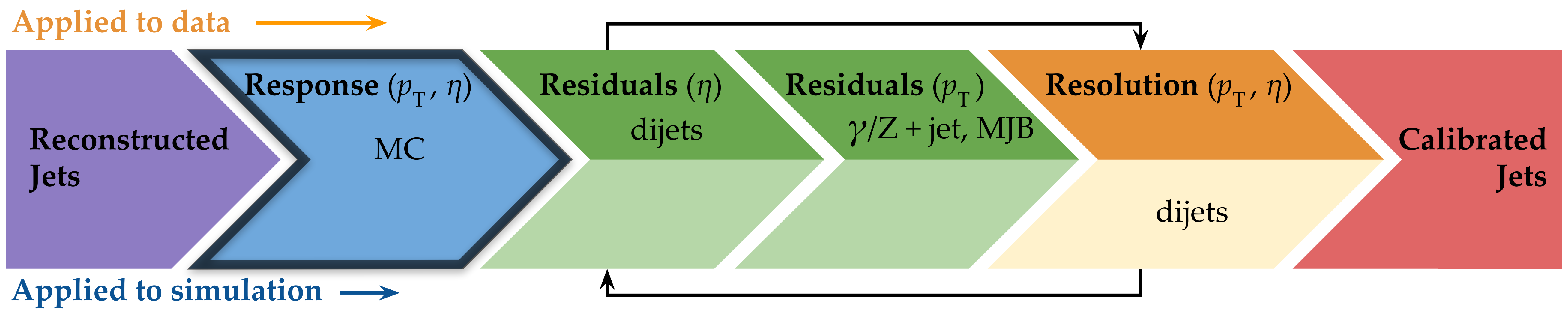
CMS-DP-2023-045



CHS

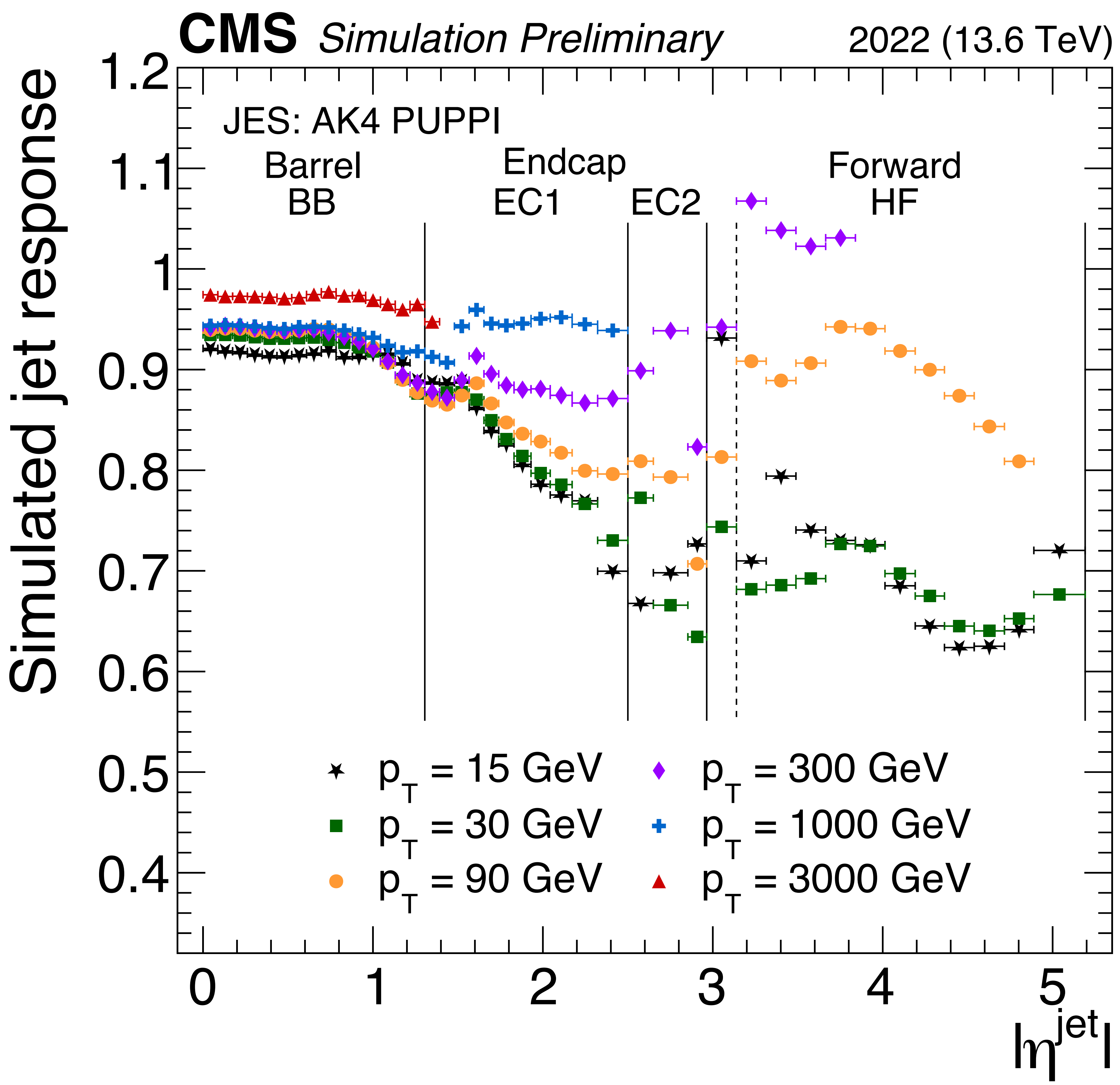
Jet calibration in CMS

L2L3 Response corrections



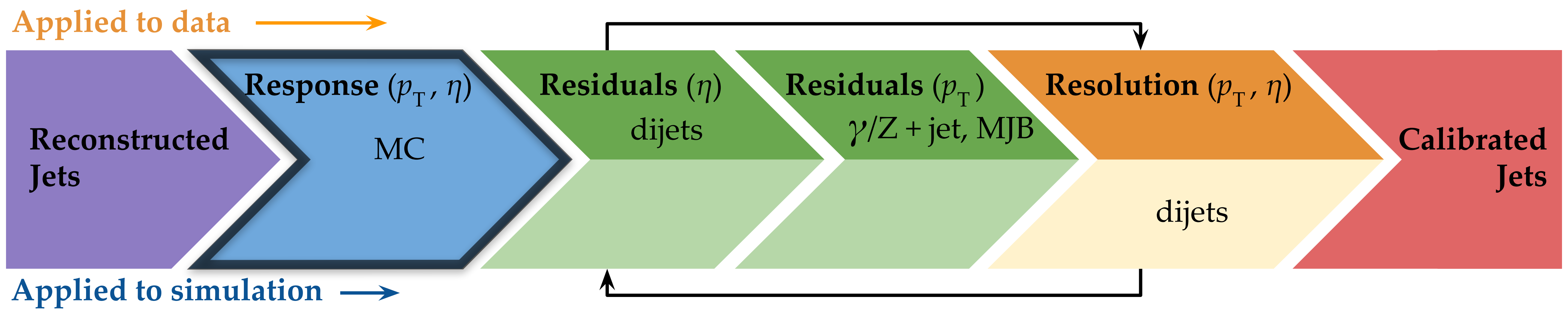
CMS-DP-2023-045

- Correction in bins of jet η and p_T correcting jet response to unity
  - Response defined as median of $p_T^{\text{reco}}/p_T^{\text{ptcl}}$
- Stable in barrel $|\eta| < 1.3$
- stronger p_T dependence in endcaps and hadron forward
- Response of JES correction closure within 0.1% to 0.4% for $|\eta| < 2.5$ and 0.8% $|\eta| > 2.5$ and $p_T > 30$ GeV



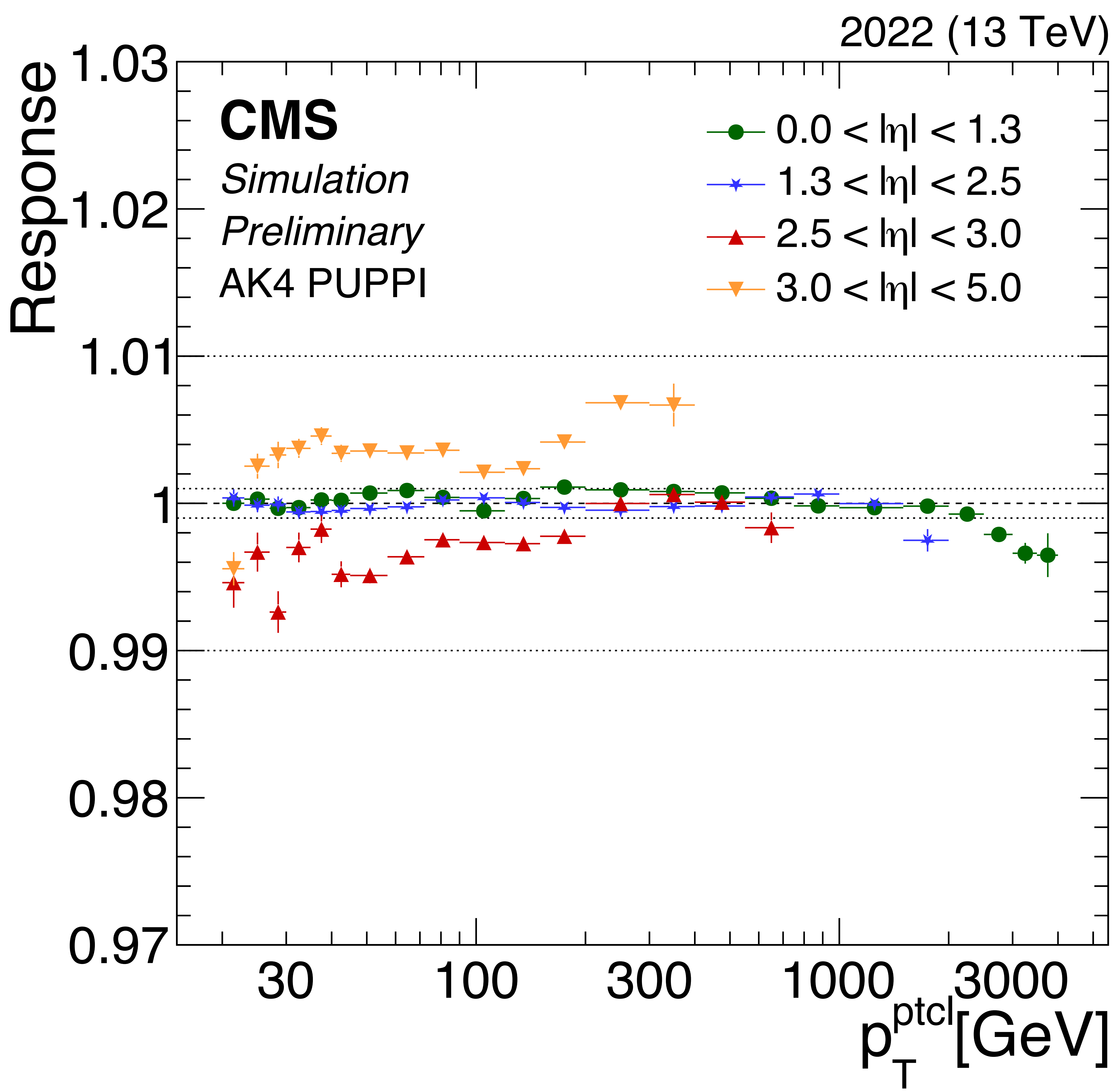
Jet calibration in CMS

L2L3 Response corrections



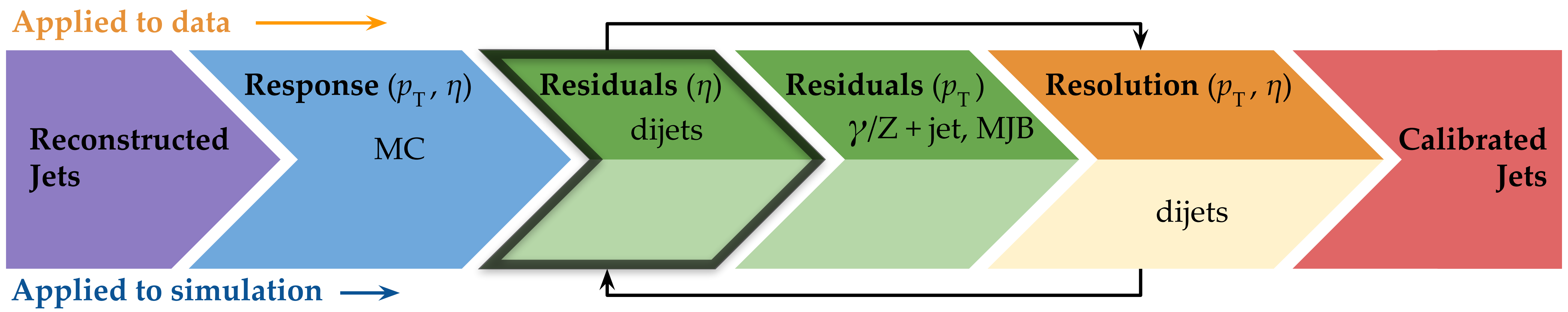
CMS-DP-2023-045

- Correction in bins of jet η and p_T correcting jet response to unity
  - Response defined as median of $p_T^{\text{reco}}/p_T^{\text{ptcl}}$
- Stable in barrel $|\eta| < 1.3$
- stronger p_T dependence in endcaps and hadron forward
- Response of JES correction closure within 0.1% to 0.4% for $|\eta| < 2.5$ and 0.8% $|\eta| > 2.5$ and $p_T > 30$ GeV

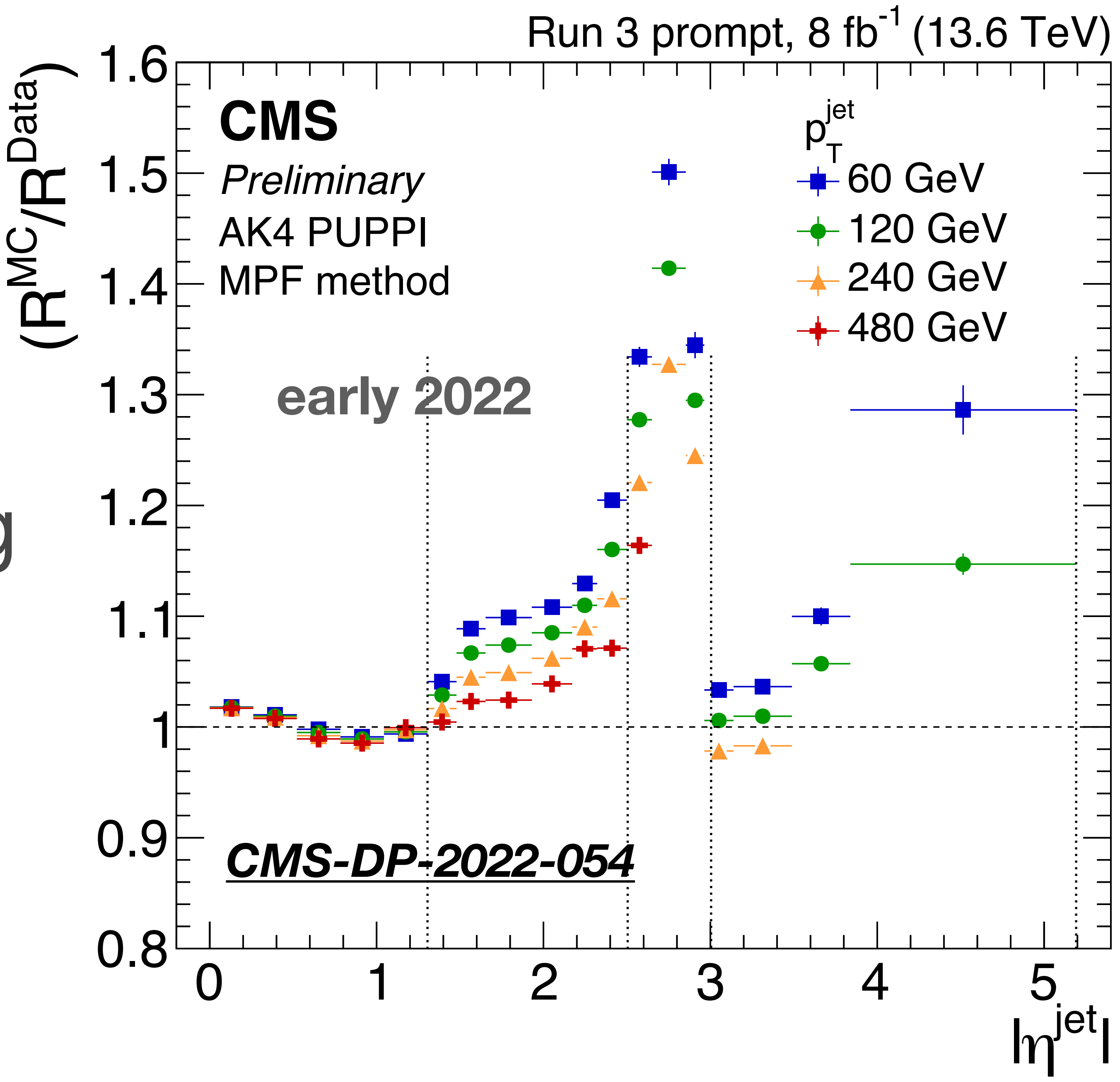


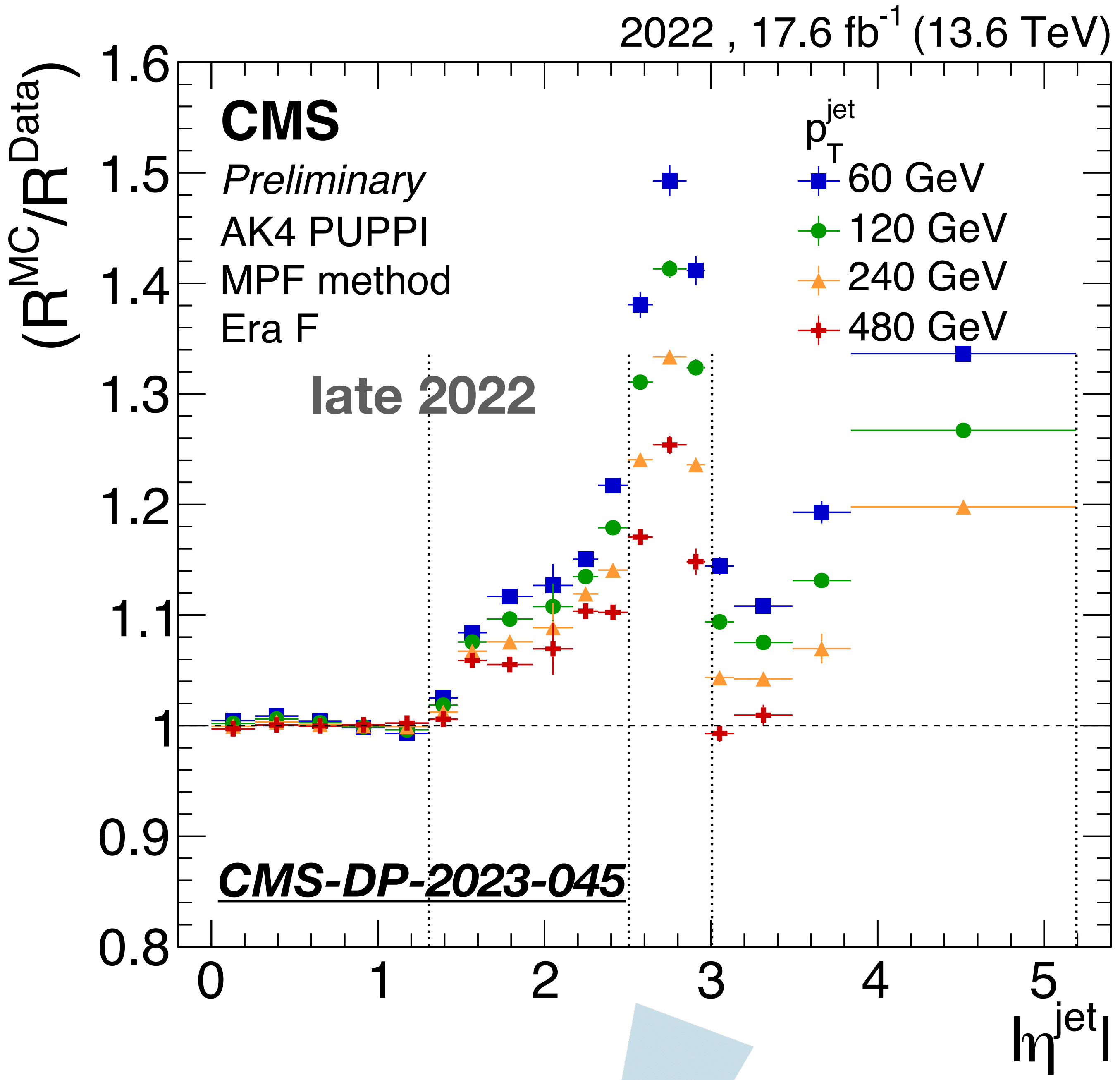
Jet calibration in CMS

L2Res η_{jet} dependent



- Residual correction in bins of jet η_{jet} and p_T^{ave}, correcting the response **relative** to the response of central jets ($|\eta| < 1.3$)
  - Derived from **dijet events** using **MPF** (missing transverse momentum projection fraction) method
  - Correction < 2% in barrel
  - In high forward endcap transition region ($2.5 < |\eta| < 3$) up to 50% and large p_T dependence
    - Likely due to missing calibration of HE

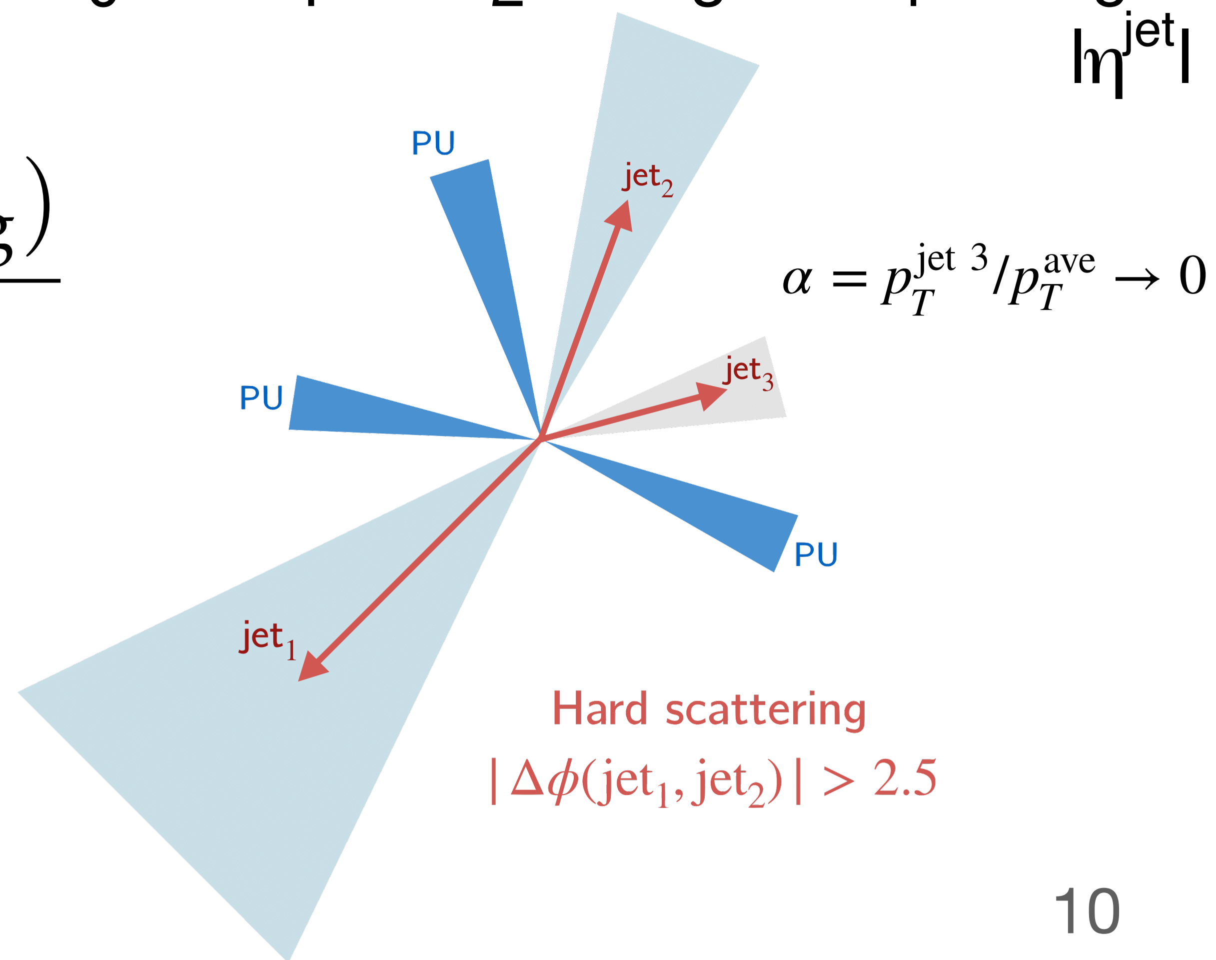




$$\mathcal{B} = \frac{\vec{p}_{\mathrm{T}}^{\,\text{miss}} \cdot (\vec{p}_{\mathrm{T,\,tag}} / p_{\mathrm{T,\,tag}})}{2 p_{\mathrm{T,\,ave}}}$$

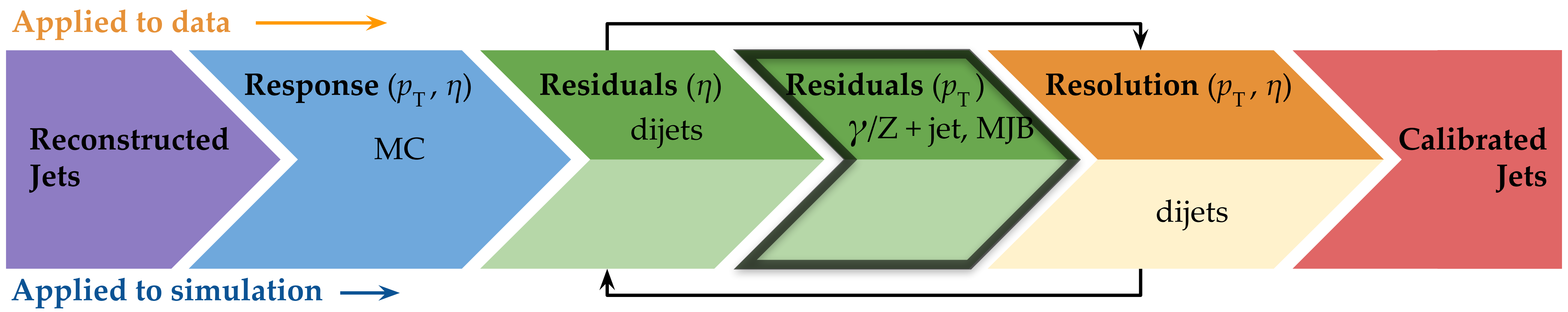
$$R_{\text{rel}}^{\text{MPF}} = \frac{1 + \langle \mathcal{B} \rangle}{1 - \langle \mathcal{B} \rangle}$$

$$p_T^{\text{ave}} = (p_T^{\text{jet 1}} + p_T^{\text{jet 2}})/2$$



Jet calibration in CMS

L3Res $p_{T,\mathrm{jet}}$ dependent

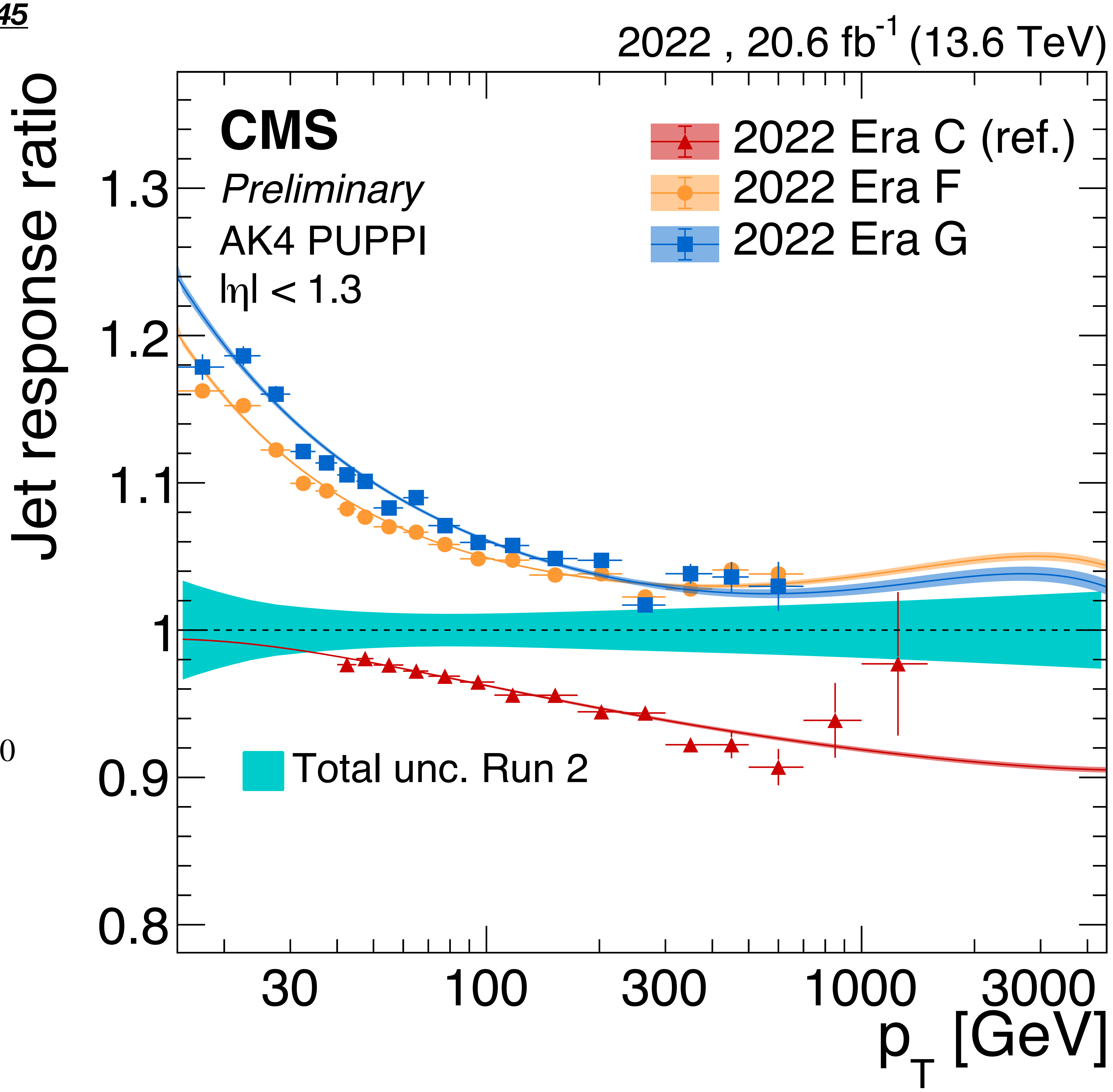


CMS-DP-2023-045

- Residual correction as a function of $p_{T,\mathrm{jet}}$
- Derived from $Z(\mu\mu)$+jets **events** after calibrating Z mass
  - using **MPF** method $$R_{\mathrm{jet,MPF}} = 1 + \frac{\vec{p}_T^{\,\mathrm{miss}} \cdot \vec{p}_{T,Z}}{(p_{T,Z})^2}$$
  - and **Direct Balance (DB)** method $$R_{\mathrm{jet,DB}} = \frac{p_{T,\mathrm{jet}}}{p_{T,Z}}$$

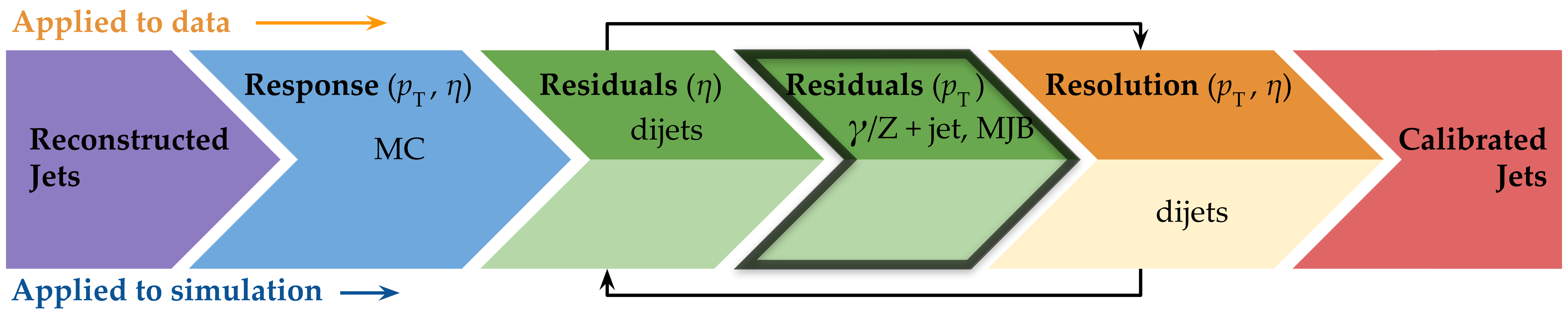
$\alpha = p_T^{\mathrm{jet}\,2}/p_{T,Z} \to 0$

- Consistent with HCAL barrel scale shift due to miscalibration: 0.90 (Era C), 1.08 (Era F), and 1.10 (Era G)

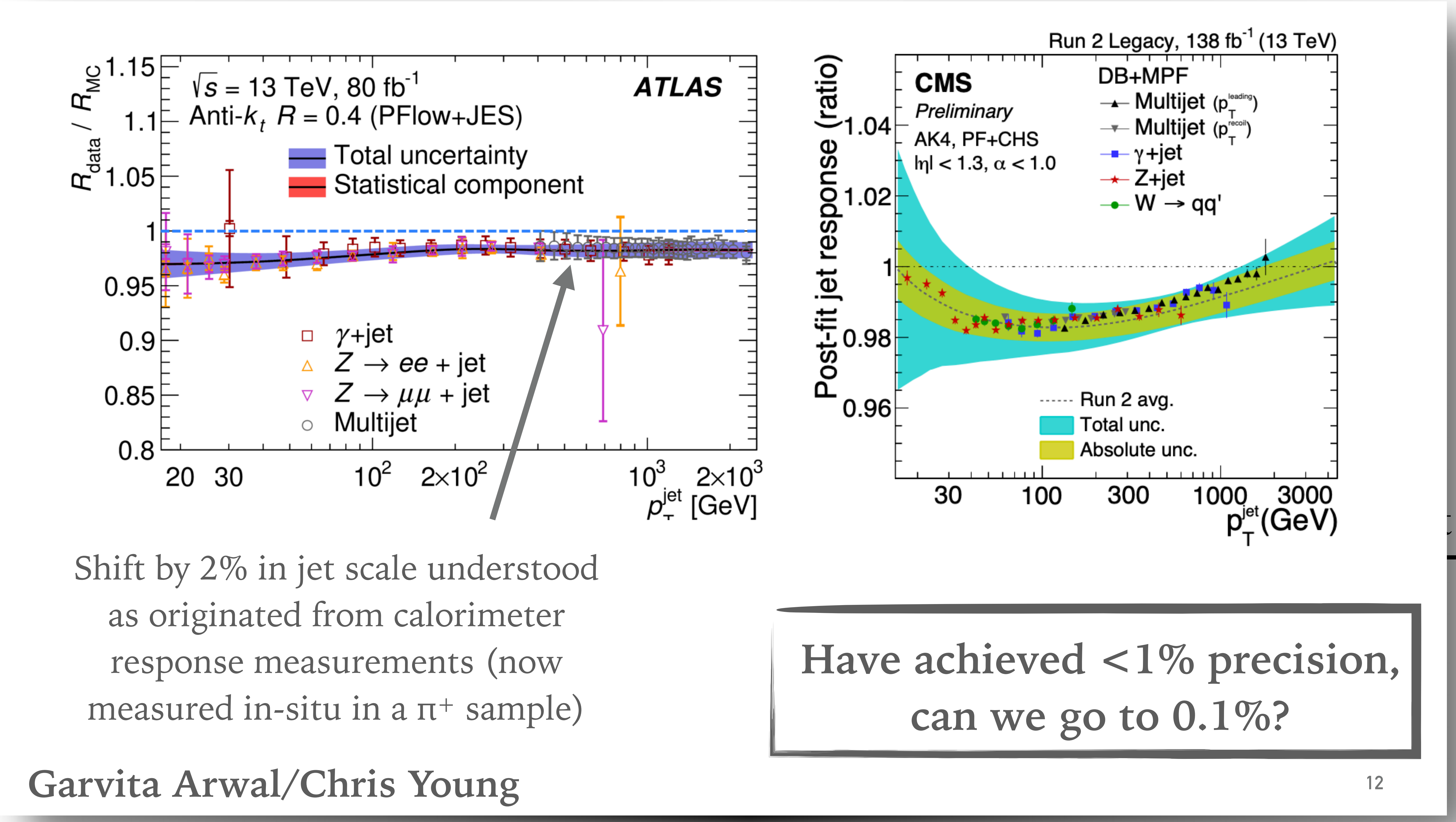


Jet calibration in CMS

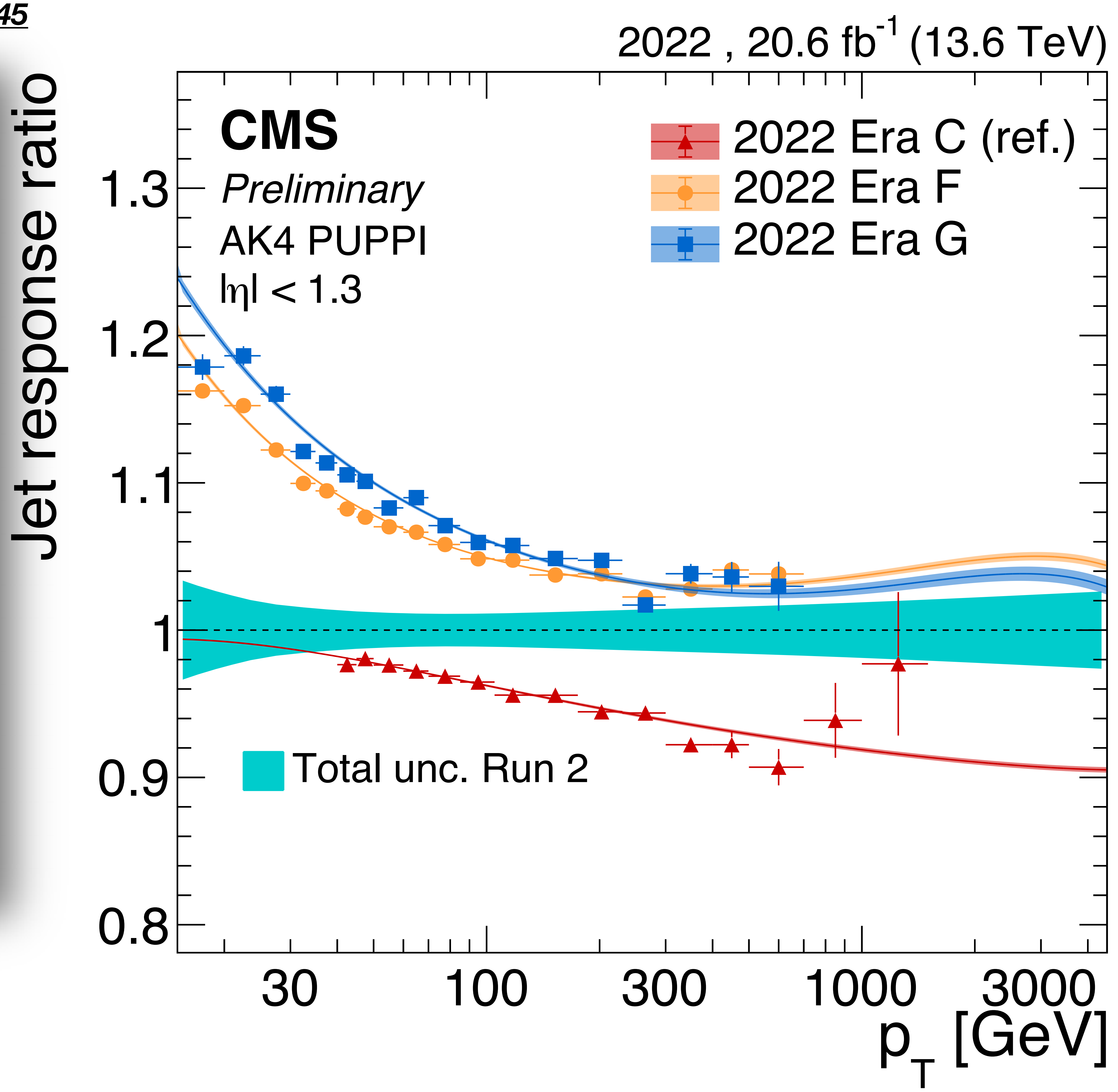
L3Res $p_{T,\text{jet}}$ dependent



Summary talk by Cristina @ BOOST22

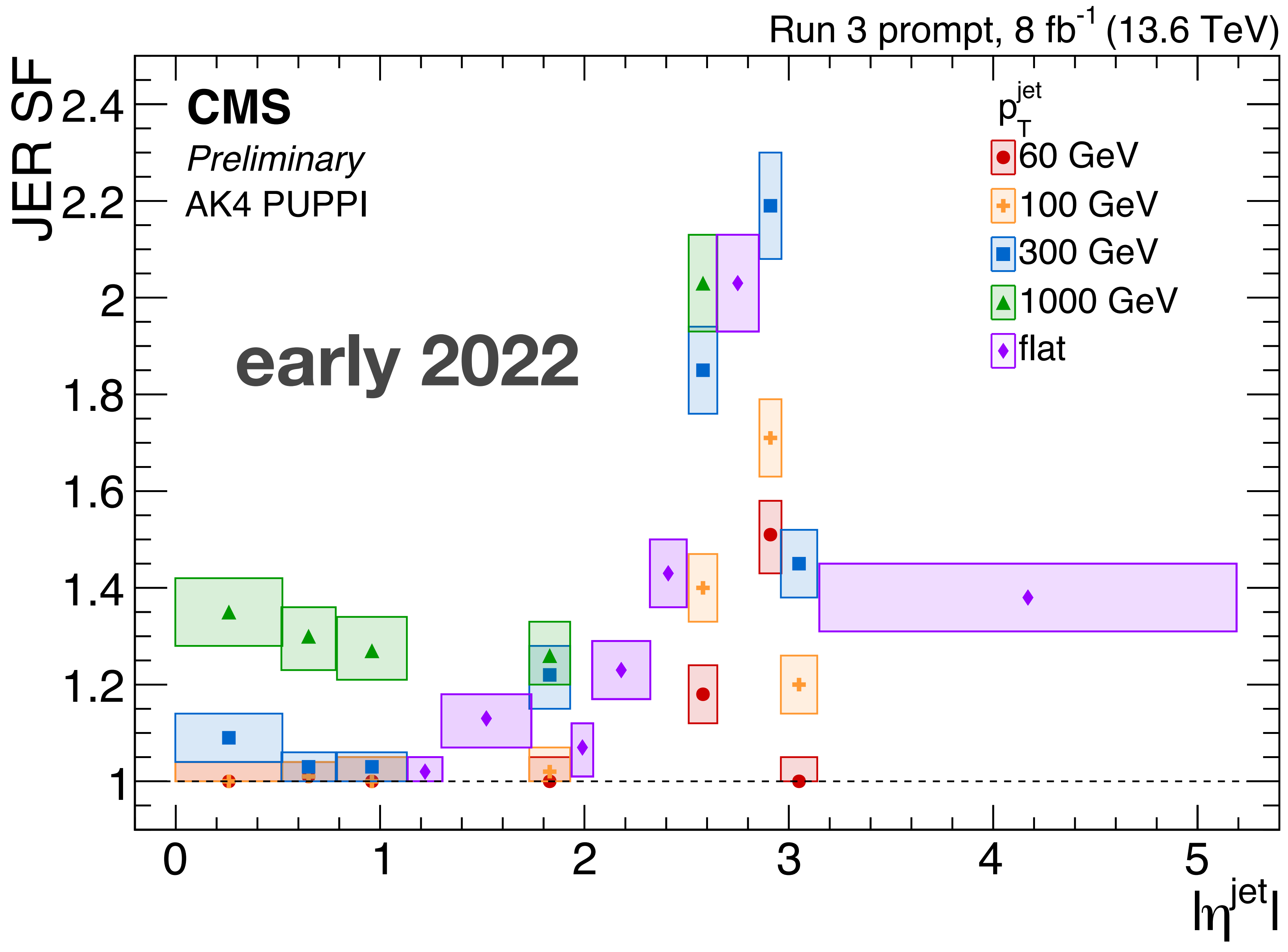


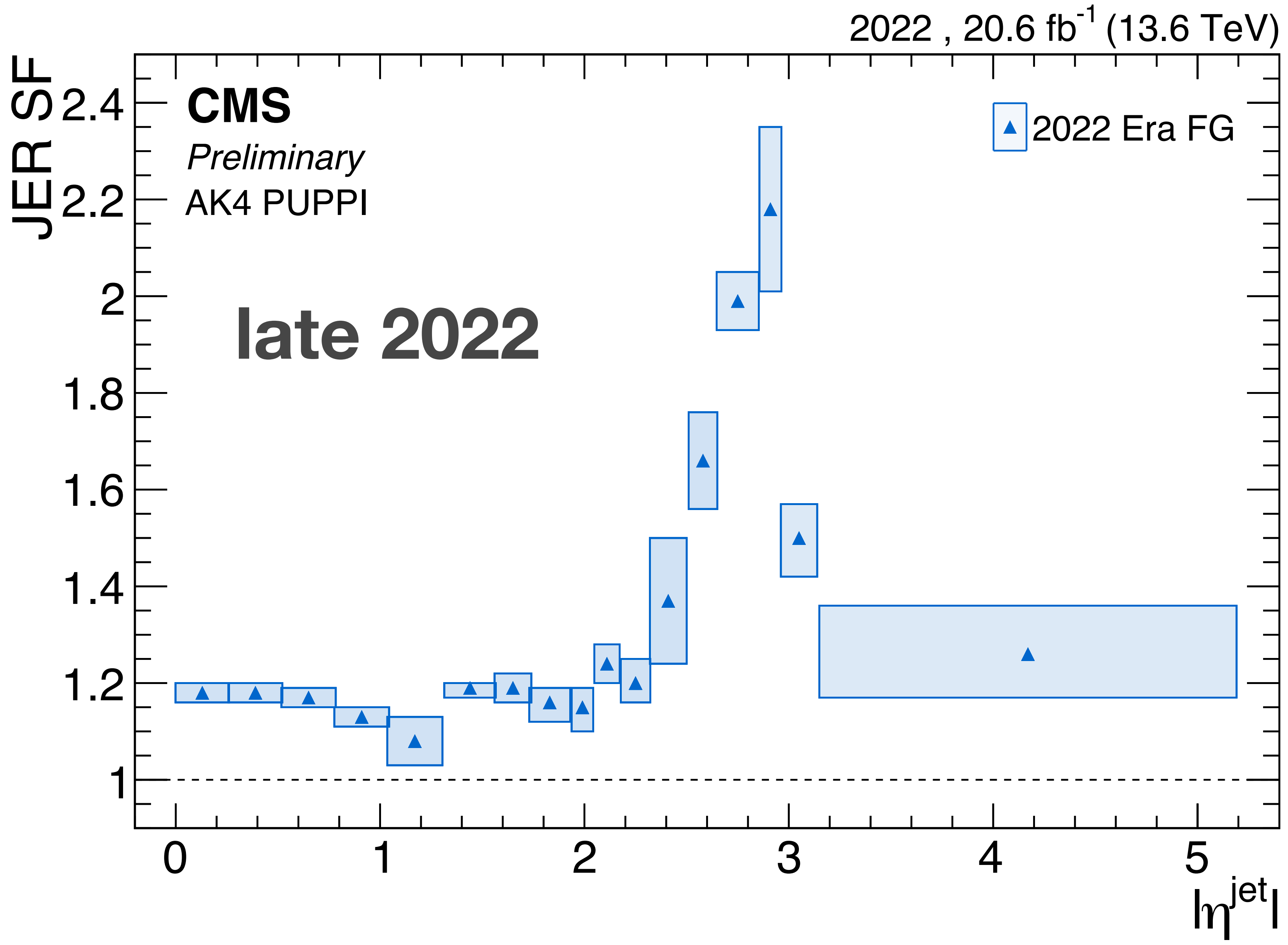
CMS-DP-2023-045



For sub percent precision, we need to calibrate the detector well

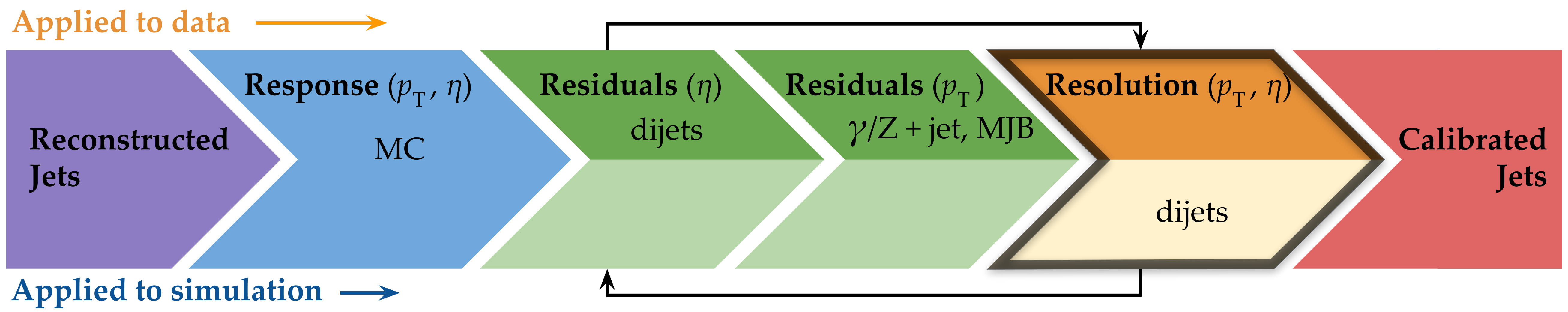
Jet calibration in CMS





CMS-DP-2022-054

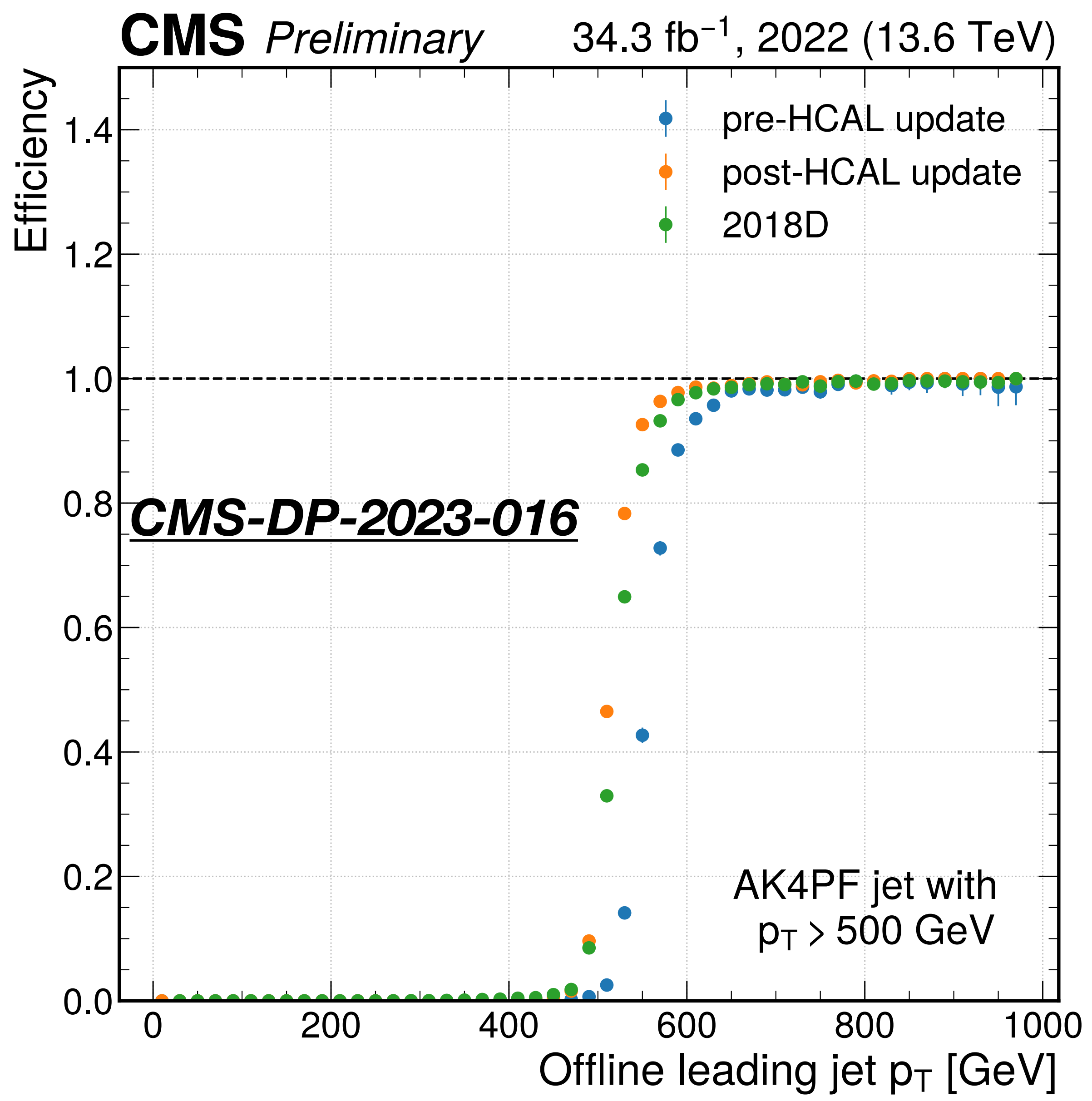
CMS-DP-2023-045

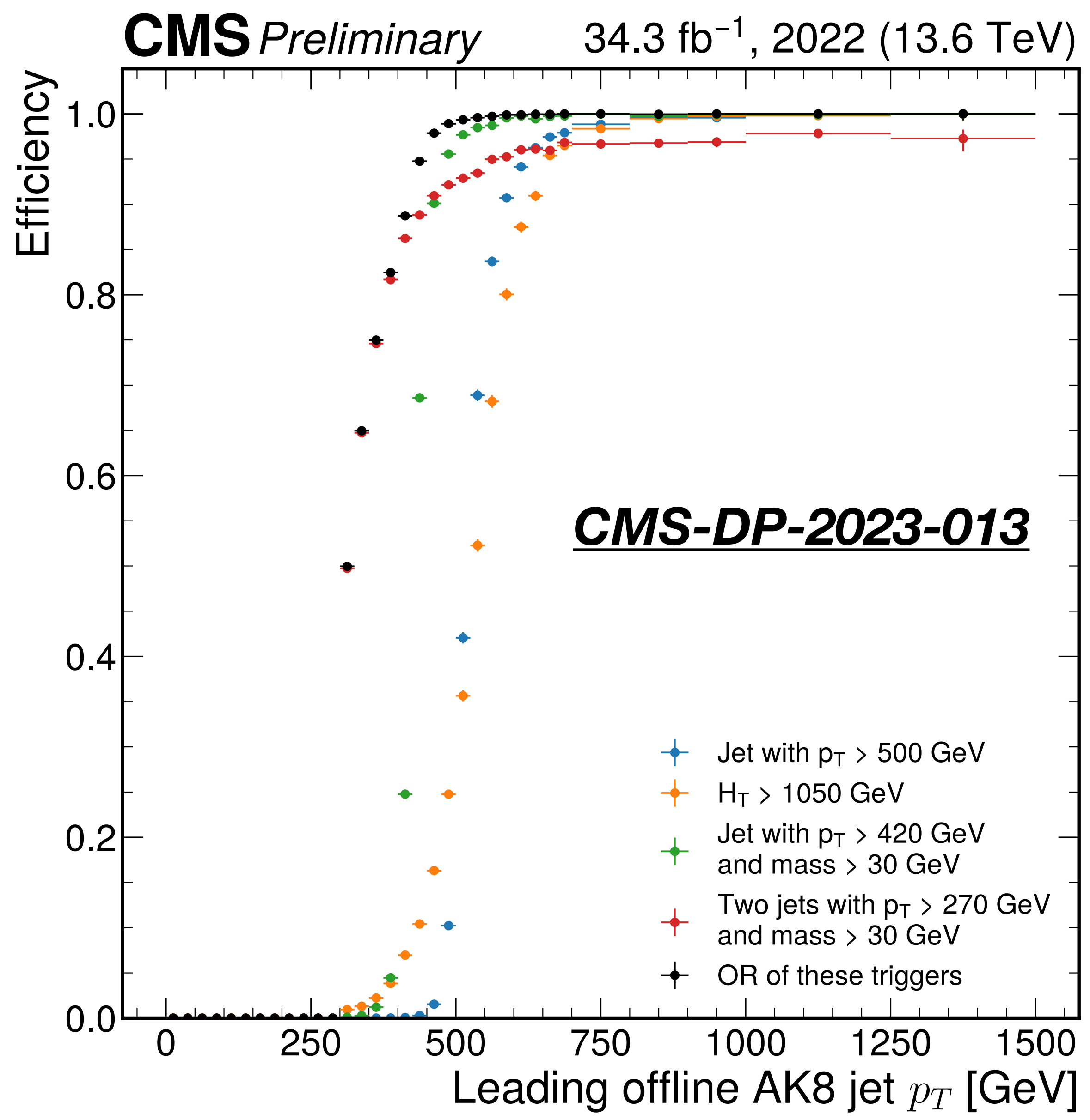


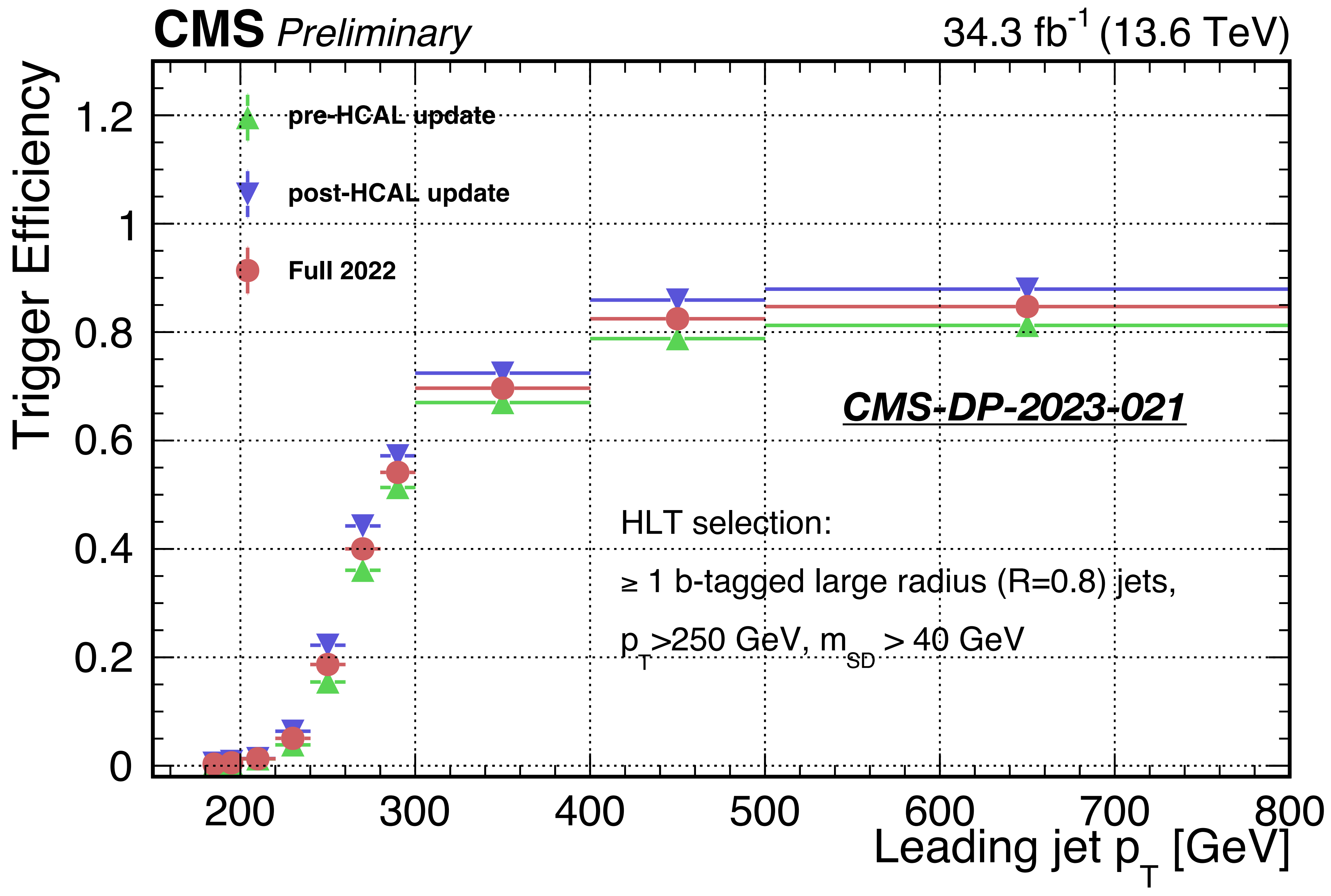
η_{jet} dependent JER SF

- Jet energy resolution correction for simulation derived in bins of $|\eta^{\text{jet}}|$
  - Measured in dijet, extrapolated to $\alpha = p_{T,\text{jet 3}}/p_T^{\text{ave}} \rightarrow 0$
  - JER SF around 1.0 to 1.4, larger in EC-HF transition $(2.5 < |\eta| < 3.0)$
  - early 2022 shows p_T dependence in some $|\eta^{\text{jet}}|$-bins

JetMET related HLT-Trigger







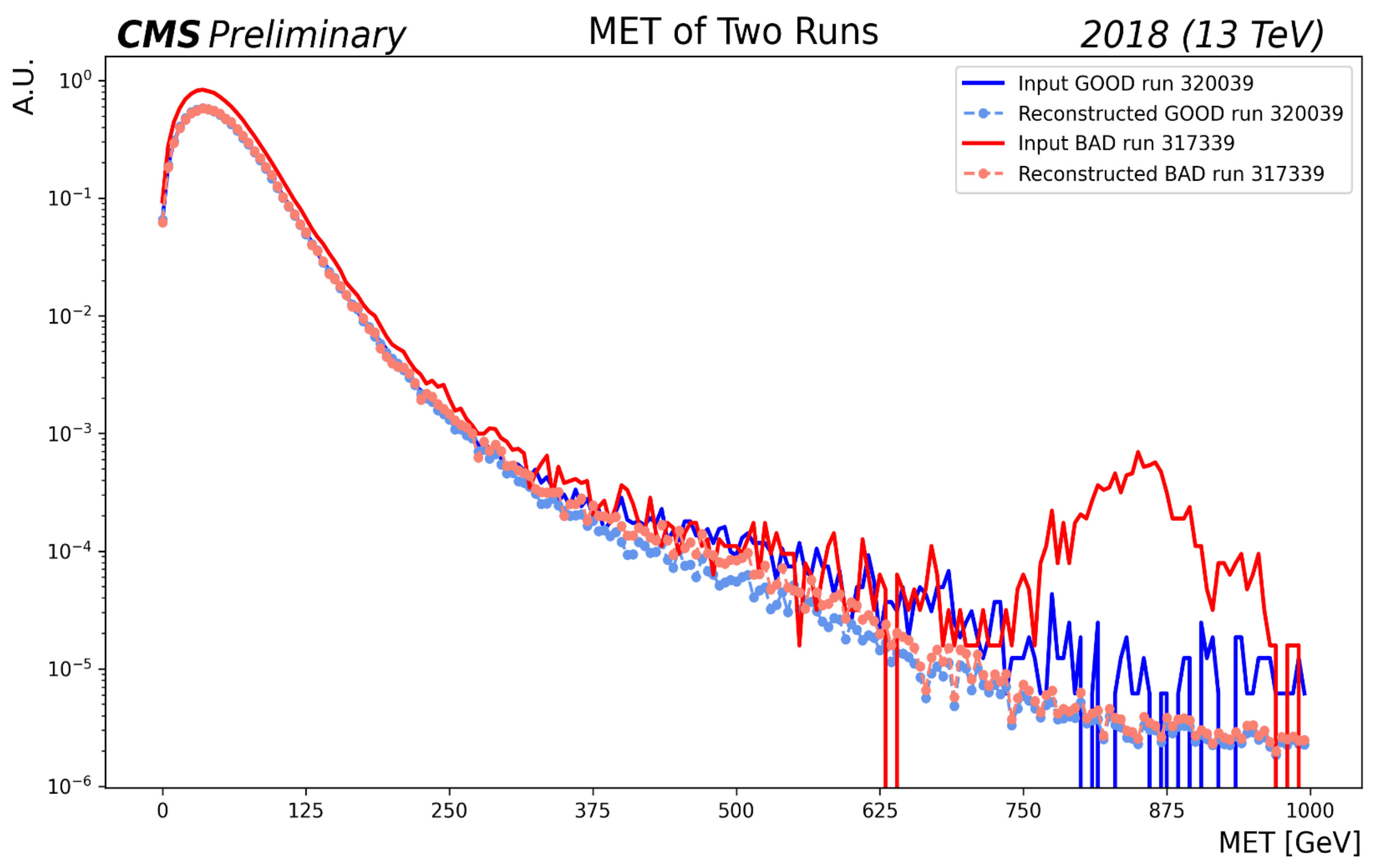
- JetMET related HLT triggers (single jet, HT, MET, soft drop mass) well behaved
  - HCAL calibration update: pre-update miscalibration corrected with residual JECs
- ParticleNet @ HLT
  - AK8 with $m_{\mathrm{SD}} > 40\ \mathrm{GeV}$, $p_T > 250\ \mathrm{GeV}$ and high $H(bb)$ vs QCD. ParticleNet score
  - Reaches up to ≈ 85% efficiency

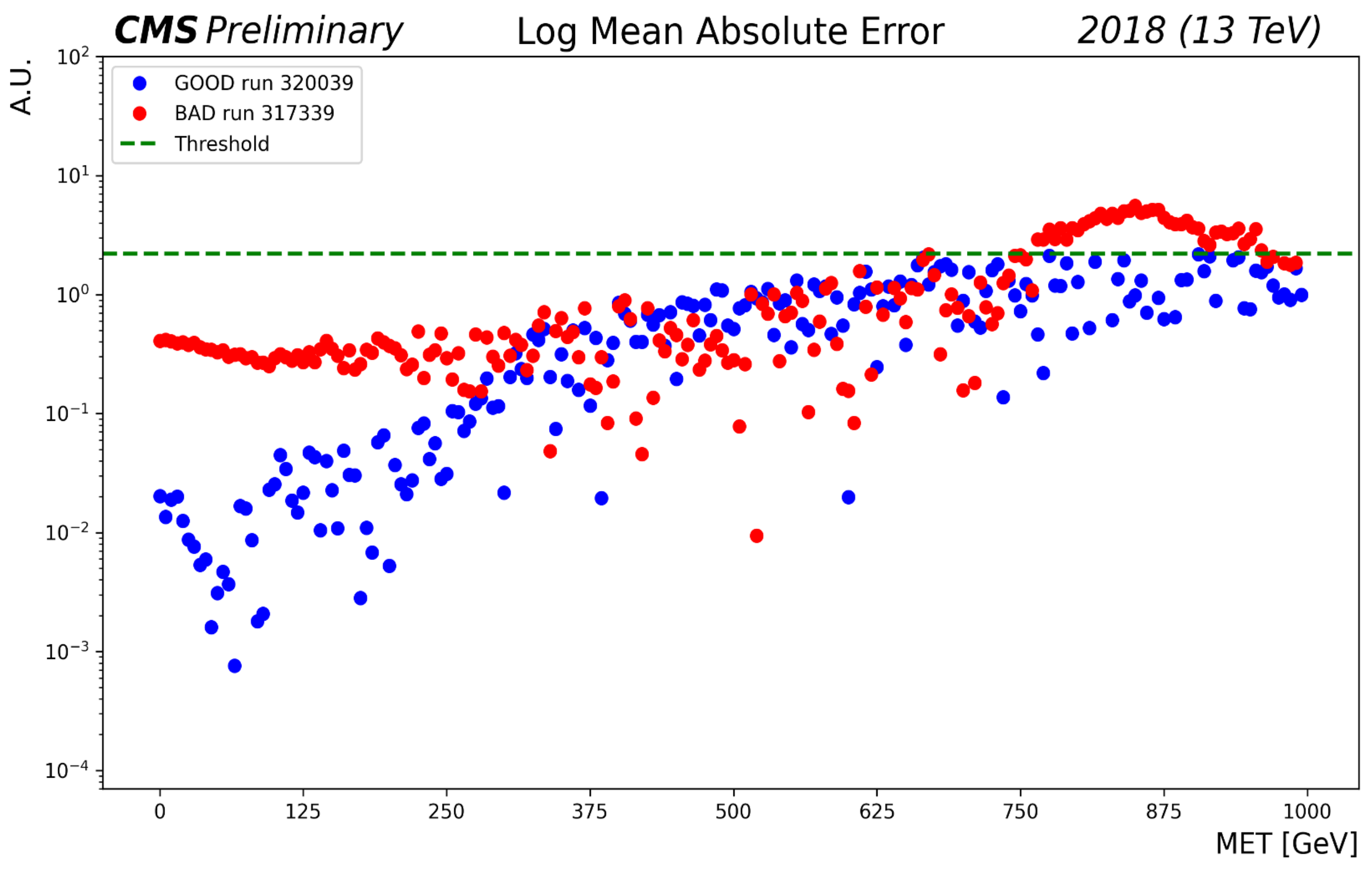
ML for Data certification

AutoEncoders

CMS-DP-2023-032

- DQM and DC online using many monitoring elements
  - Time intensive and prone to human error
- AutoEncoders (AC) with JetMET monitoring elements as input
  - non-anomalous runs as input
  - Threshold on reconstruction loss as metric for GOOD/BAD runs
- For anomalous runs look at a per lumi section (LS) basis
  - recover BAD runs partially by rejecting only anomalous LS



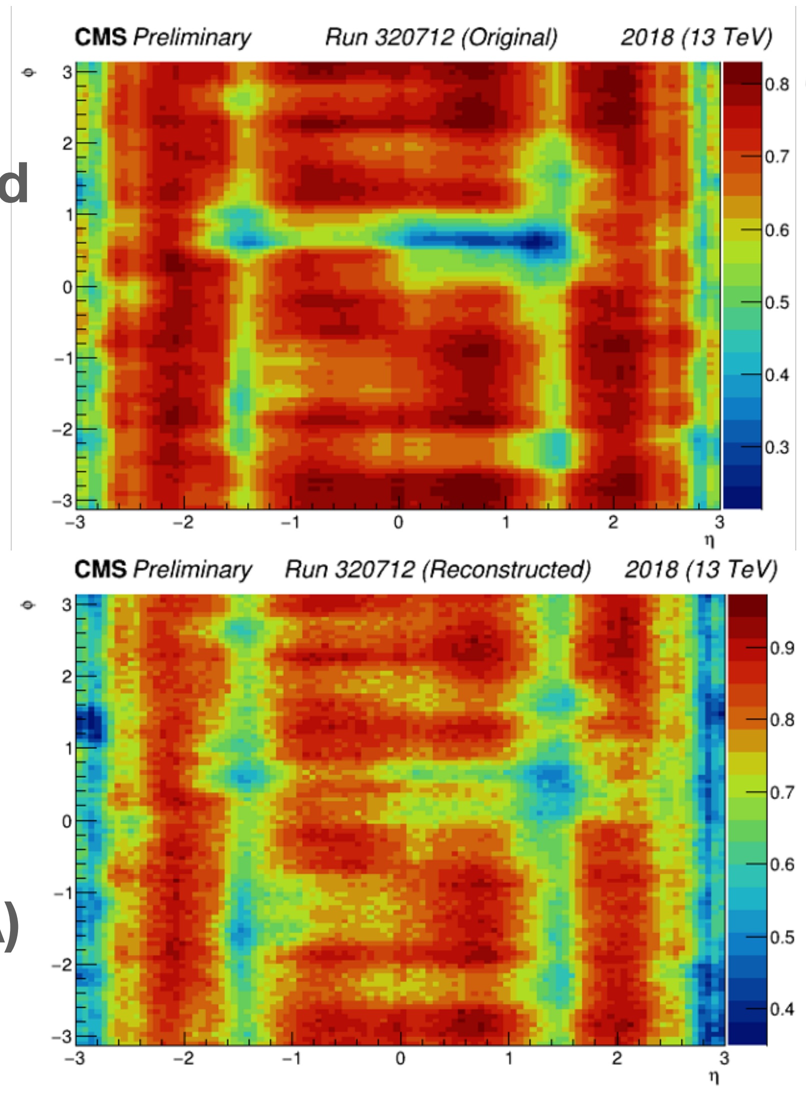


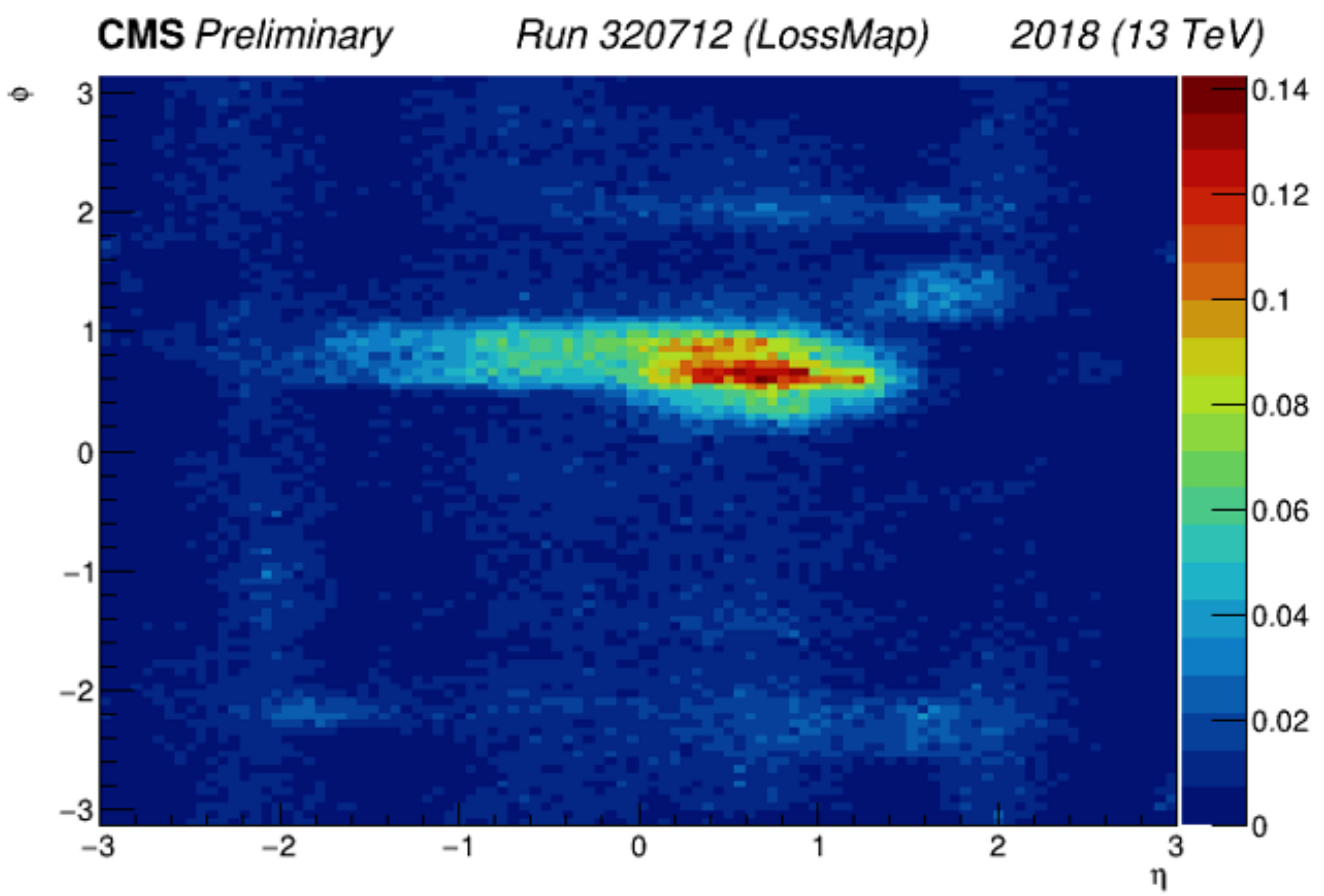
Performance on par with supervised networks

2018 Era D

Measured

AC reconstruction (trained on Era A)





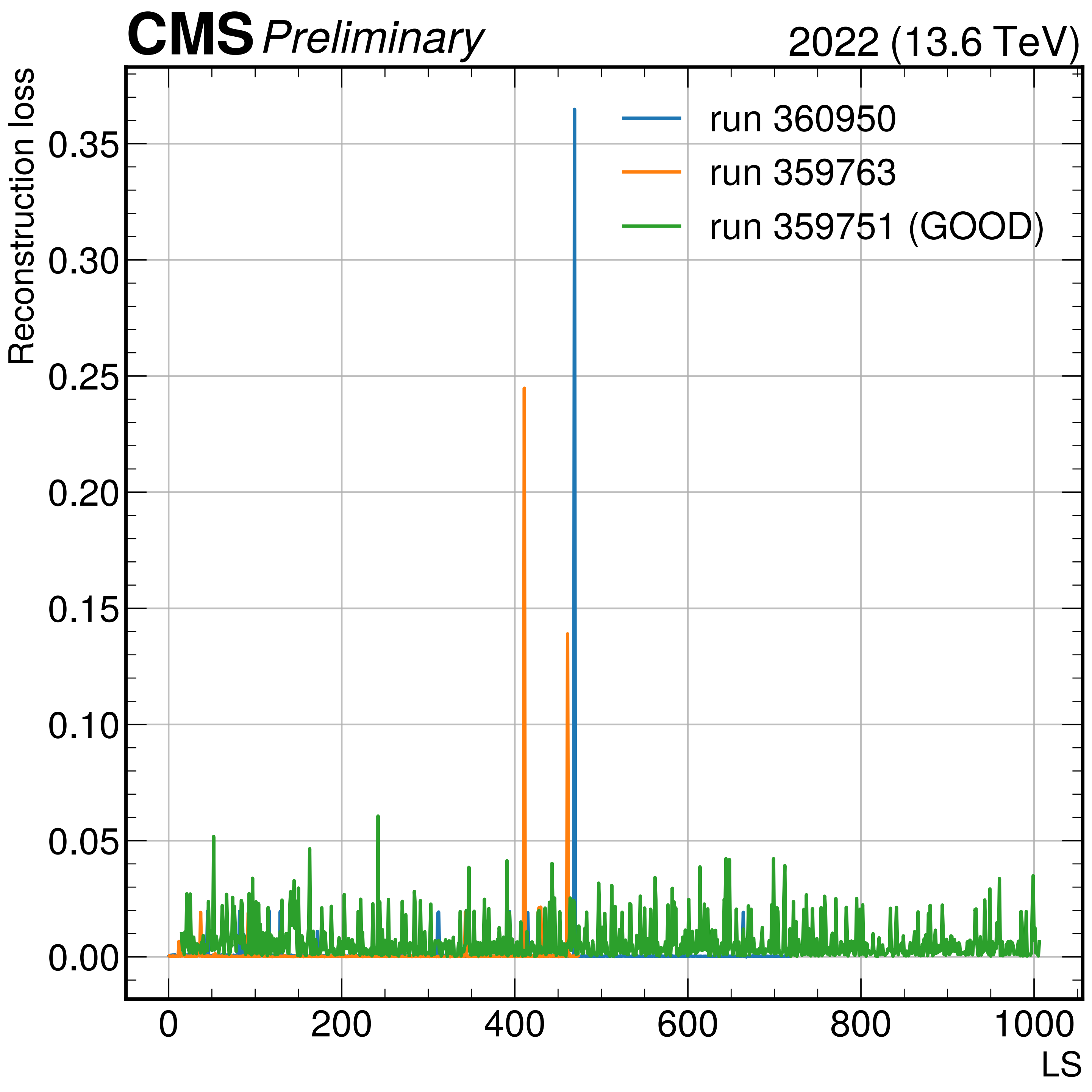
Multilayer perceptron trained with 2D histograms can **also** differentiate between data taking eras

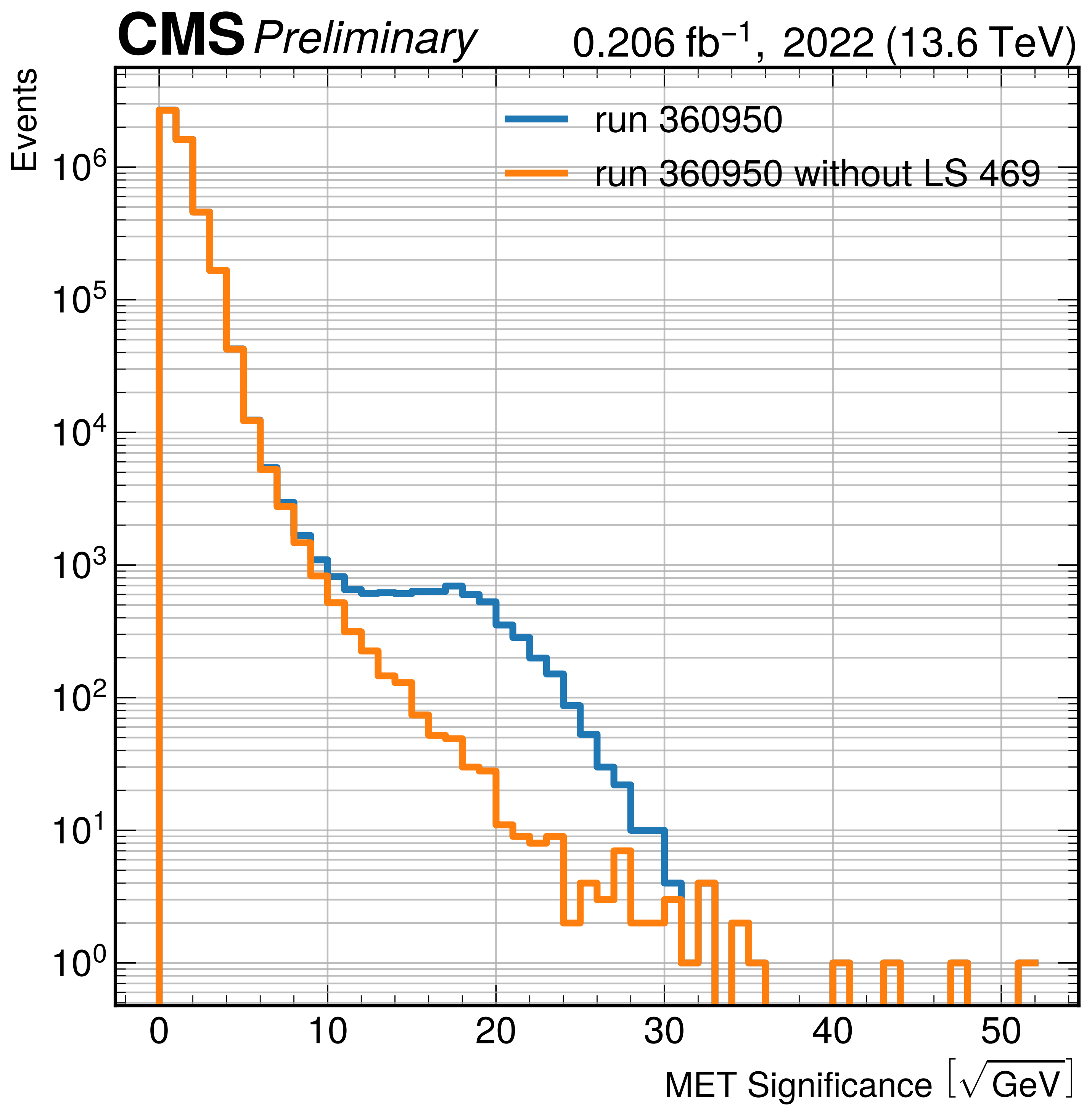
ML for Data certification

AutoEncoders

CMS-DP-2023-010

- DQM and DC online using many monitoring elements
  - Time intensive and prone to human error
- AutoEncoders (AC) with JetMET monitoring elements as input
  - non-anomalous runs as input
  - Threshold on reconstruction loss as metric for GOOD/BAD runs
- For anomalous runs look at a per lumi section (LS) basis
  - recover BAD runs partially by rejecting only anomalous LS





Summary

- PUPPI performs well in Run 3
- Jet energy calibration corrects for miscalibration of detector
  - will improve as detector is calibrated further
- Jet&MET related trigger well behaved in Run 3
  - New ParticleNET H(bb) vs. QCD tagger
- Study of Autoencoders and other ML trained on Jet&MET monitoring elements for use in Data Certification

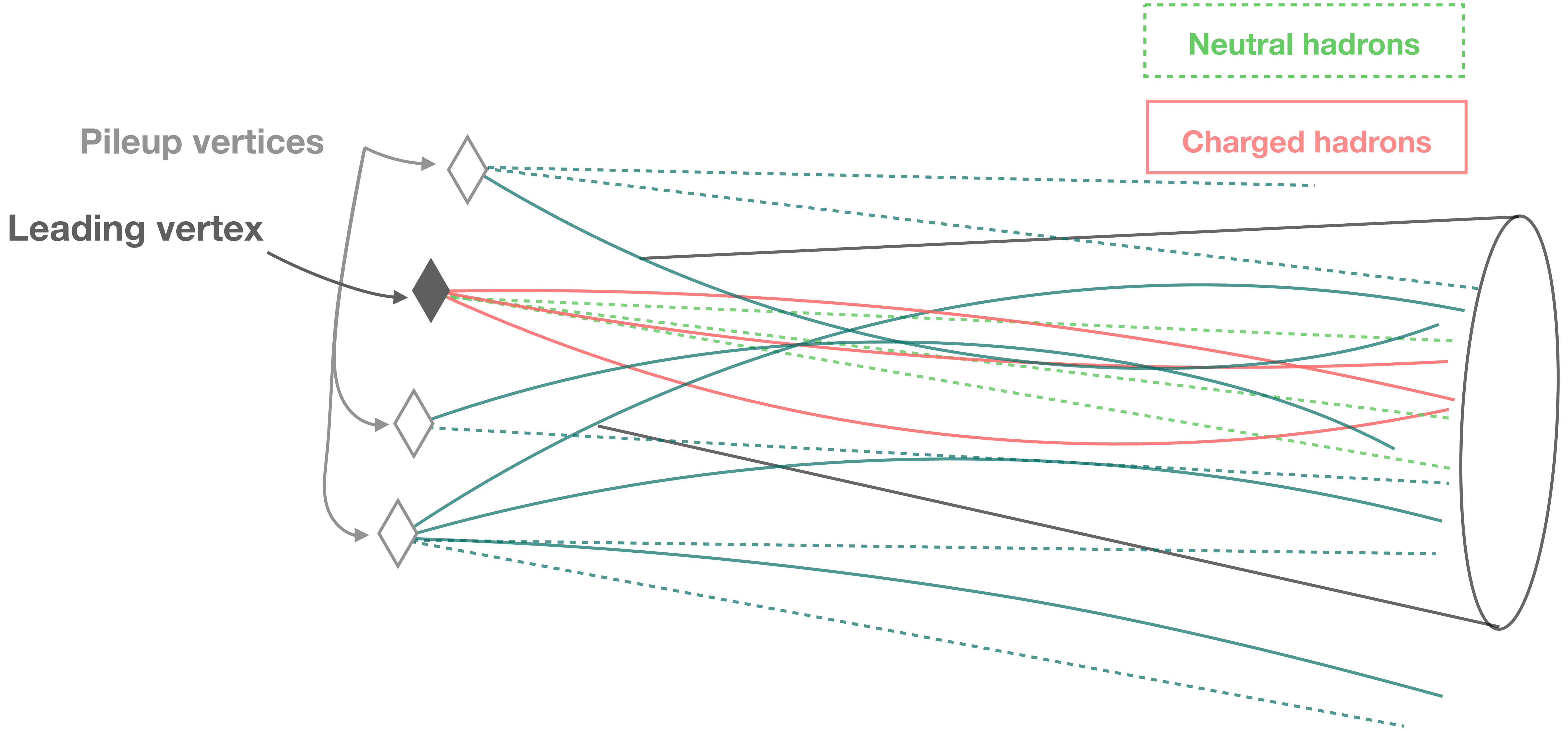
Additional Material

References

CMS-DP-2022-054	Jet Energy Scale and Resolution Measurements Using Prompt Run 3 Data Collected by CMS in the First Months of 2022 at 13.6 TeV
CMS-DP-2023-045	Jet Energy Scale and Resolution Measurements Using Prompt Run 3 Data Collected by CMS in the Last Months of 2022 at 13.6 TeV
CMS-DP-2023-032	Machine Learning Techniques for JetMET Data Certification of the CMS Detector
CMS-DP-2023-010	An AutoEncoder-based Anomaly Detection tool with a per-LS granularity
CMS-DP-2023-016	Performance of jets and missing transverse momentum reconstruction at the High Level Trigger using Run 3 data from the CMS Experiment at CERN
CMS-DP-2023-013	Performance of Soft Drop Mass Jet High Level Trigger at $\sqrt{s}$ = 13.6 TeV in Run 3
CMS-DP-2023-021	Performance of the ParticleNet tagger on small and large-radius jets at High Level Trigger in Run 3
CMS-DP-2023-012	A first look at early 2022 proton-proton collisions at $\sqrt{s}$=13.6 TeV for heavy-flavor jet tagging
CMS-DP-2022-063	Performance of JetMET high level trigger algorithms in the CMS experiment using proton-proton collisions data at $\sqrt{s}$=13 TeV during Run-2
CMS-DP-2023-007	Performances of Muons, Jets and MET Level 1 trigger algorithms in Run 3

Pileup mitigation in Run 3

- Pileup interactions from same or other bunch crossing overlay the hard scattering
  - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far



CHS

- Remove charged hadrons not originating from leading vertex
- Only in tracker coverage and only on charged component of pileup
  - Additional treatment on jets clustered from CHS particle collection needed: pileup jet ID and offset correction
- Used in Run 2

PUPPI

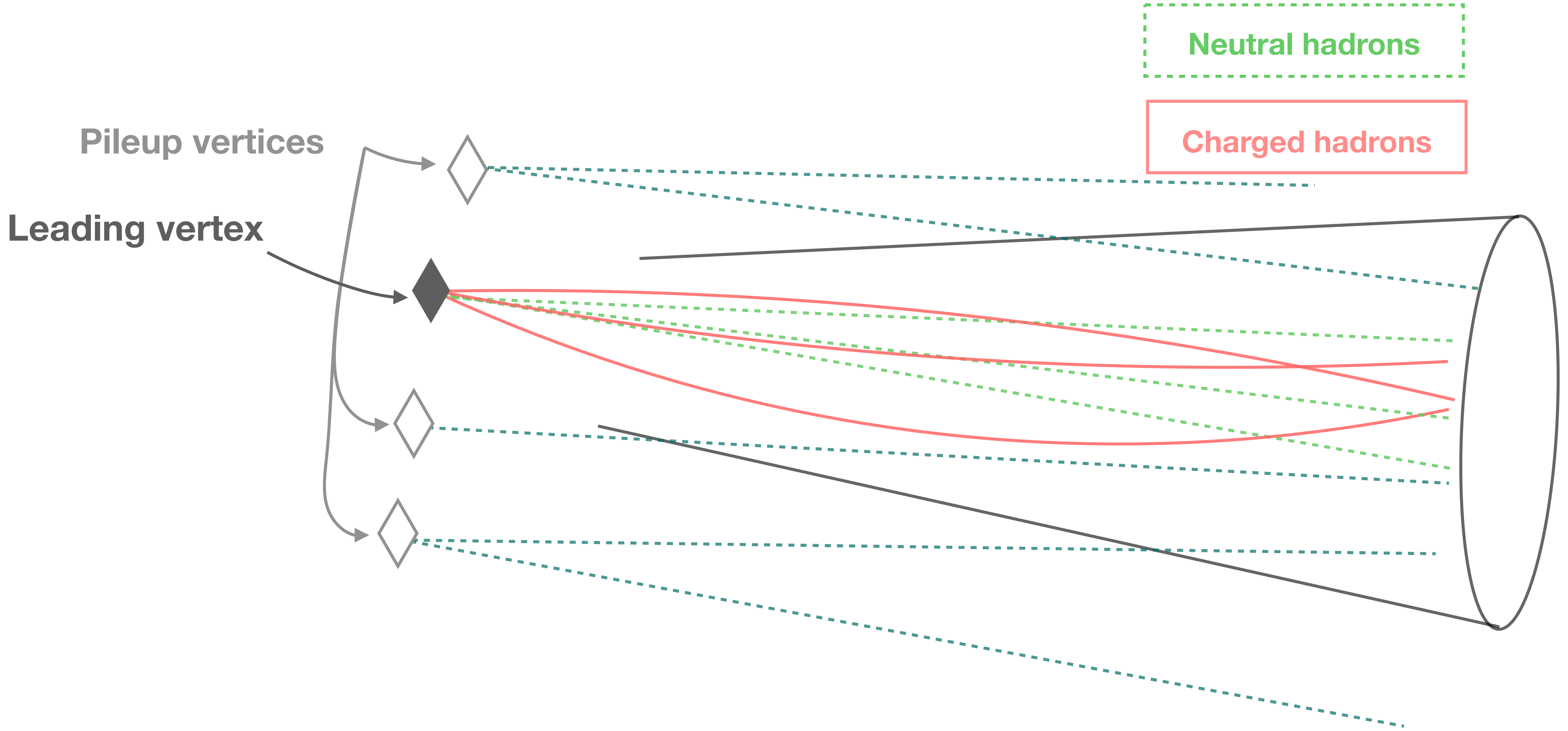
- Treating both charged and neutral particles
  - neutral particles get assigned a weight based on α_i → scales four-momenta
  - how far is it from charged particle originating from LV ≈ how likely is it they originate from LV or PV

$$\alpha_i = \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2$$

$$w_i = F_{\chi^2, \mathrm{ndf}=1} \left(\frac{(\alpha_i - \bar{\alpha}_{\mathrm{PU}}) \,|\, \alpha_i - \bar{\alpha}_{\mathrm{PU}} \,|}{(\alpha_{\mathrm{PU}}^{\mathrm{RMS}})^2} \right)$$

Pileup mitigation in Run 3

- Pileup interactions from same or other bunch crossing overlay the hard scattering
  - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far



CHS

- Remove charged hadrons not originating from leading vertex
- Only in tracker coverage and only on charged component of pileup
  - Additional treatment on jets clustered from CHS particle collection needed: pileup jet ID and offset correction
- Used in Run 2

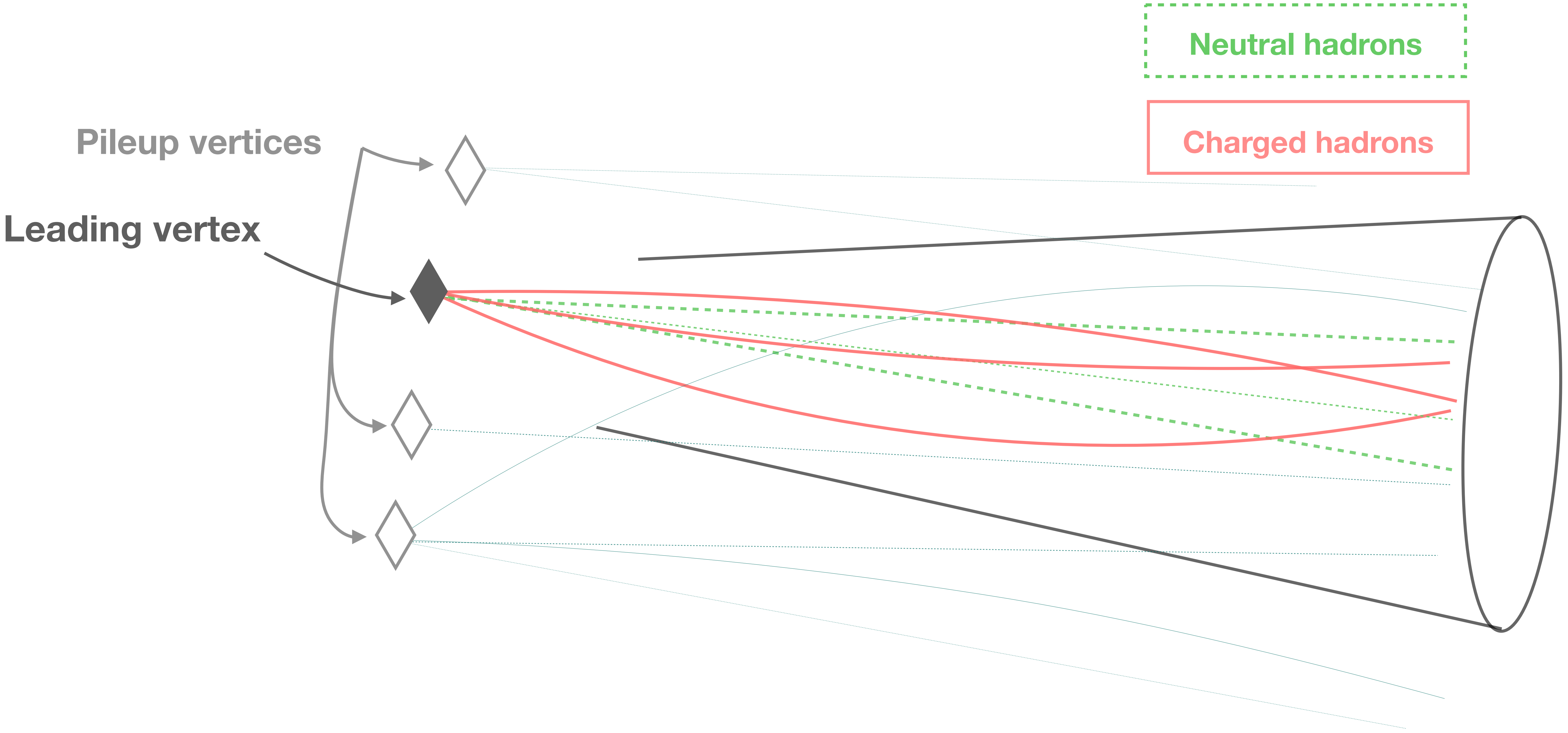
PUPPI

- Treating both charged and neutral particles
  - neutral particles get assigned a weight based on α_i → scales four-momenta
  - how far is it from charged particle originating from LV ≈ how likely is it they originate from LV or PV

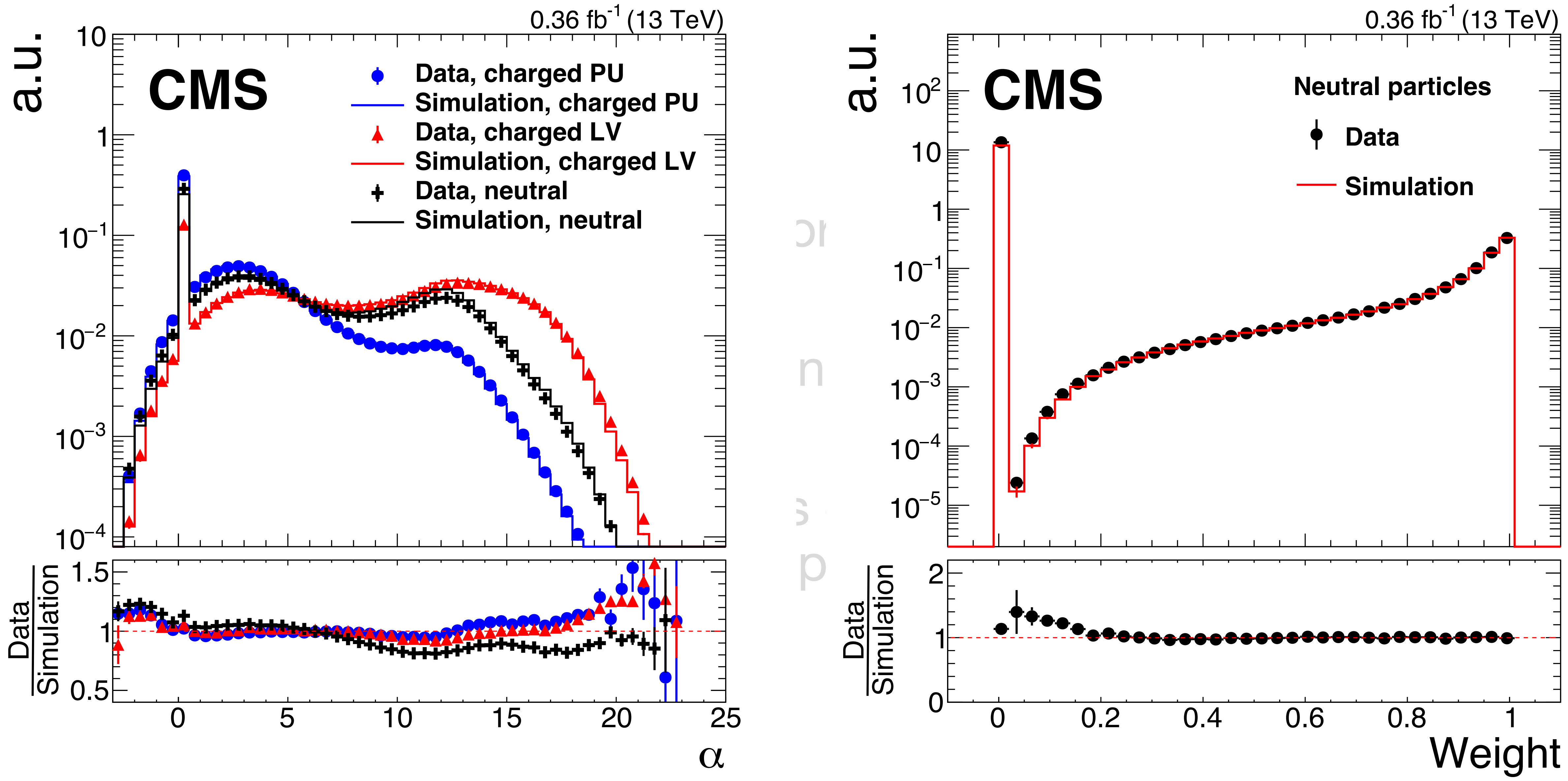
$$\alpha_i = \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \qquad w_i = F_{\chi^2, \mathrm{ndf}=1} \left(\frac{(\alpha_i - \bar{\alpha}_{\mathrm{PU}}) \, | \alpha_i - \bar{\alpha}_{\mathrm{PU}} |}{(\alpha_{\mathrm{PU}}^{\mathrm{RMS}})^2} \right)$$

Pileup mitigation in Run 3

- Pileup interactions from same or other bunch crossing overlay the hard scattering
  - Charged and neutral hadrons originating from pileup vertices end up in jet
- Need to cope with a mean number of interactions per bunch crossing of $\mu = 80$ in Run 3 so far



JINST 15 (2020) 09, P09018

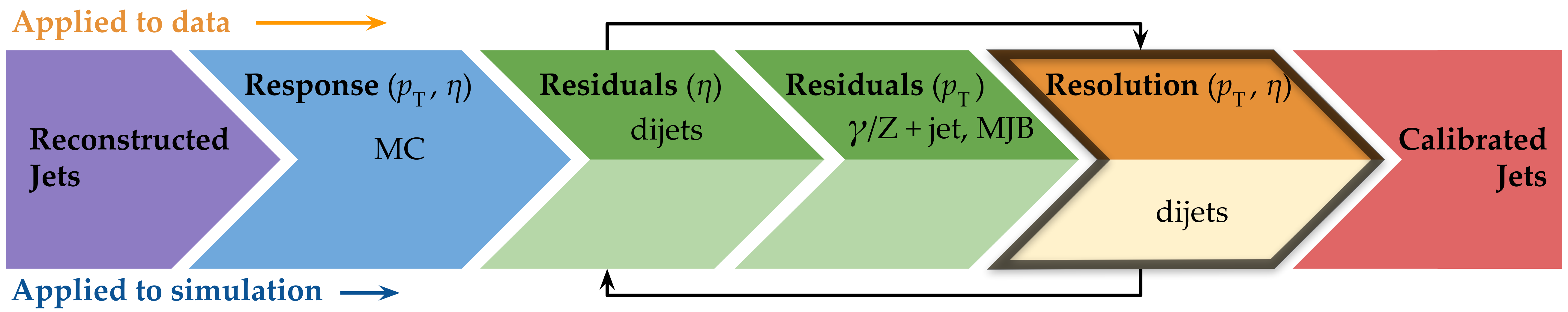


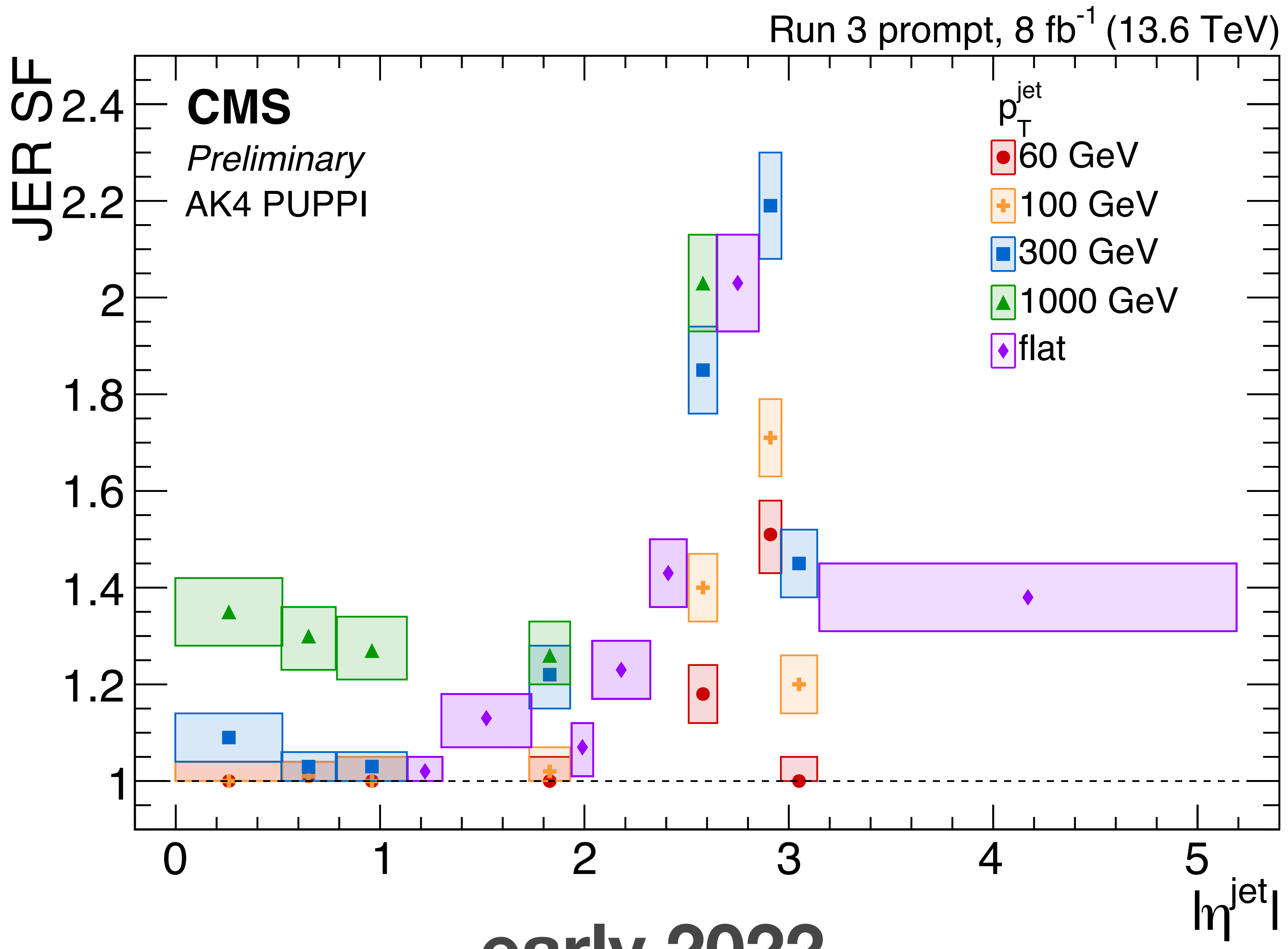
PUPPI

- Treating both charged and neutral particles
  - neutral particles get assigned a weight based on α_i → scales four-momenta
  - how far is it from charged particle originating from LV ≈ how likely is it they originate from LV or PV

$$\alpha_i = \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \qquad w_i = F_{\chi^2, \mathrm{ndf}=1} \left(\frac{(\alpha_i - \bar{\alpha}_{\mathrm{PU}})\,|\alpha_i - \bar{\alpha}_{\mathrm{PU}}|}{(\alpha_{\mathrm{PU}}^{\mathrm{RMS}})^2} \right)$$

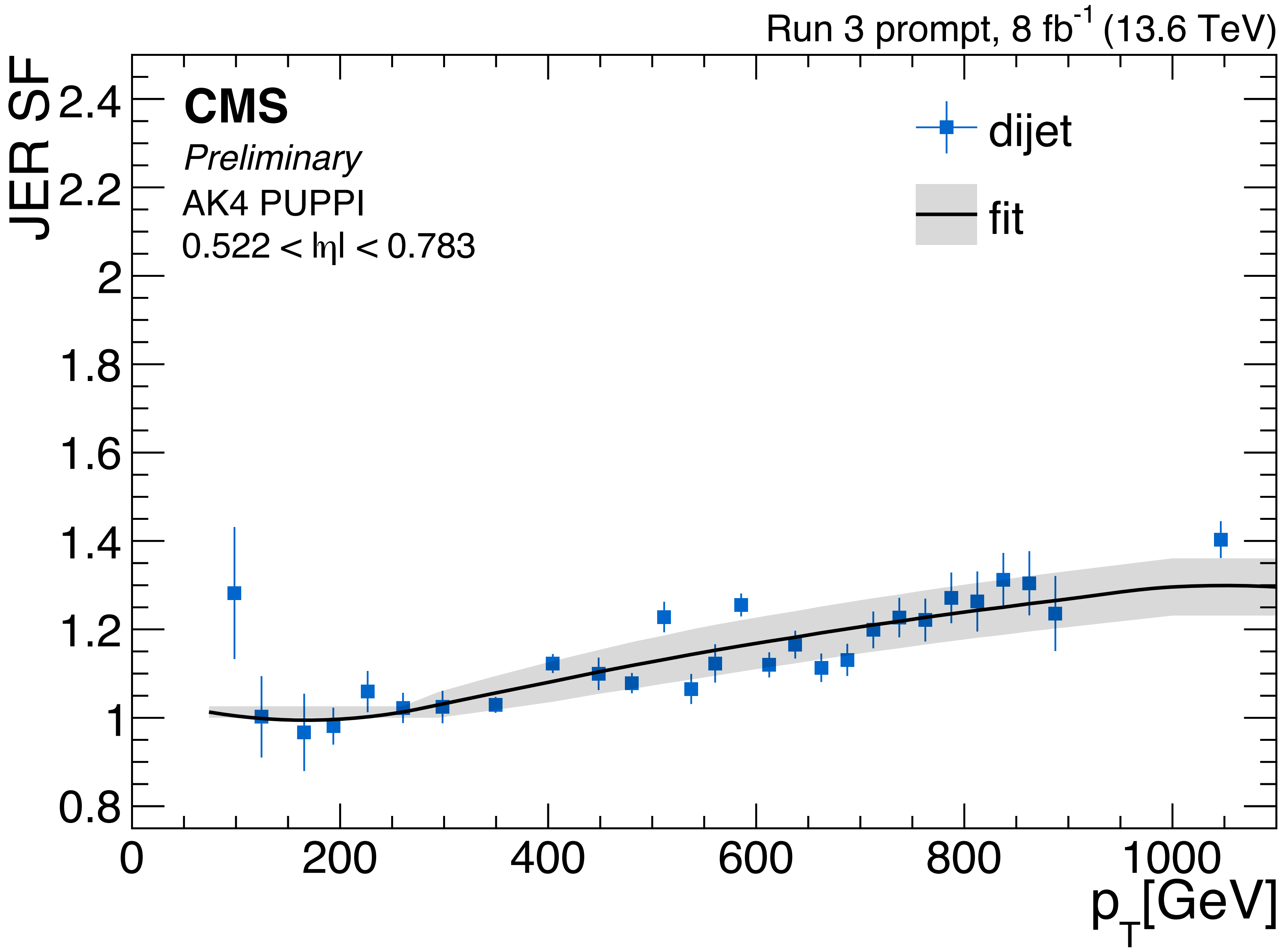
Jet calibration in CMS





CMS-DP-2022-054

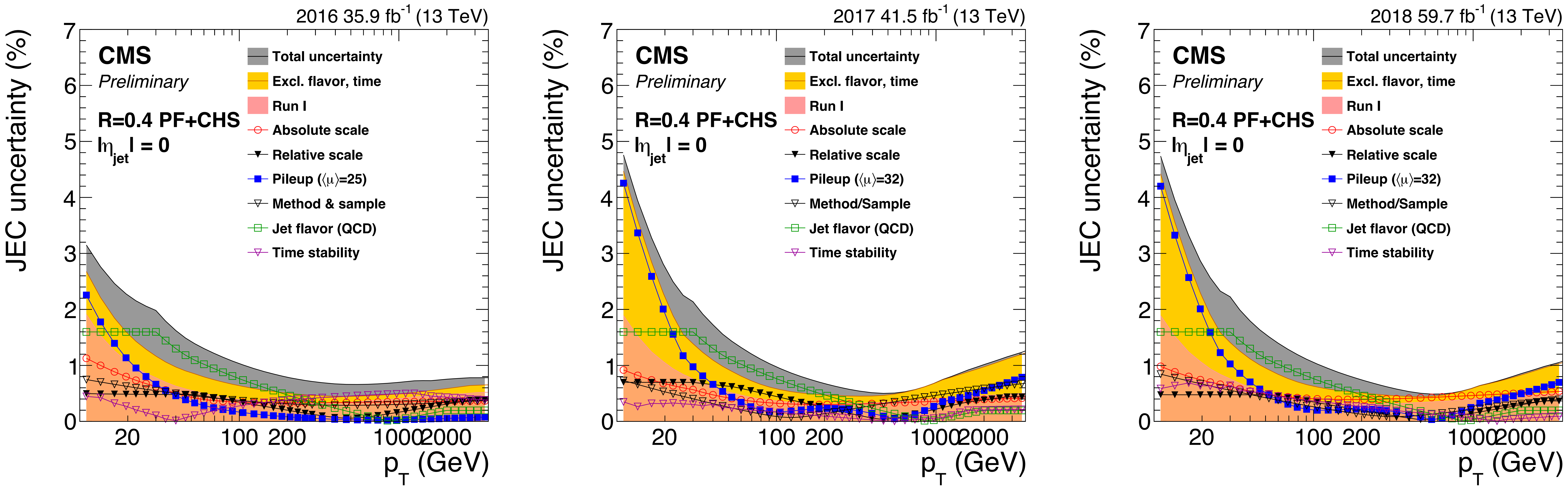
early 2022



$p_{T,\mathrm{jet}}$ dependent JER SF

- Jet energy resolution correction for simulation derived in bins of $|\eta^{\mathrm{jet}}|$
  - Measured in dijet, extrapolated to $\alpha = p_{T,\mathrm{jet}\ 3}/p_T^{\mathrm{ave}} \to 0$
  - JER SF around 1.0 to 1.4, larger in EC-HF transition $(2.5 < |\eta| < 3.0)$
  - early 2022 shows p_T dependence in some $|\eta^{\mathrm{jet}}|$-bins

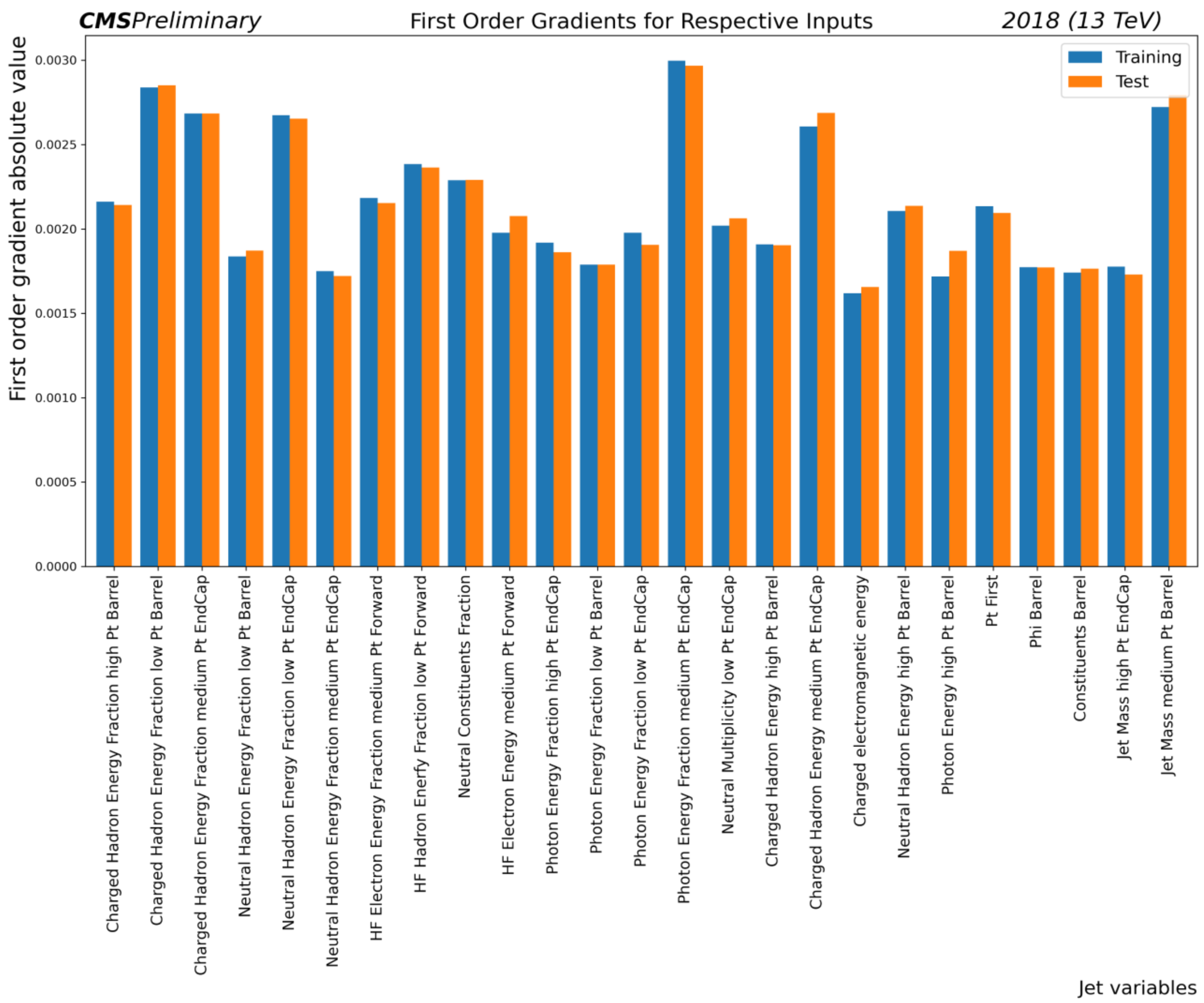
JEC Uncertainty Run II

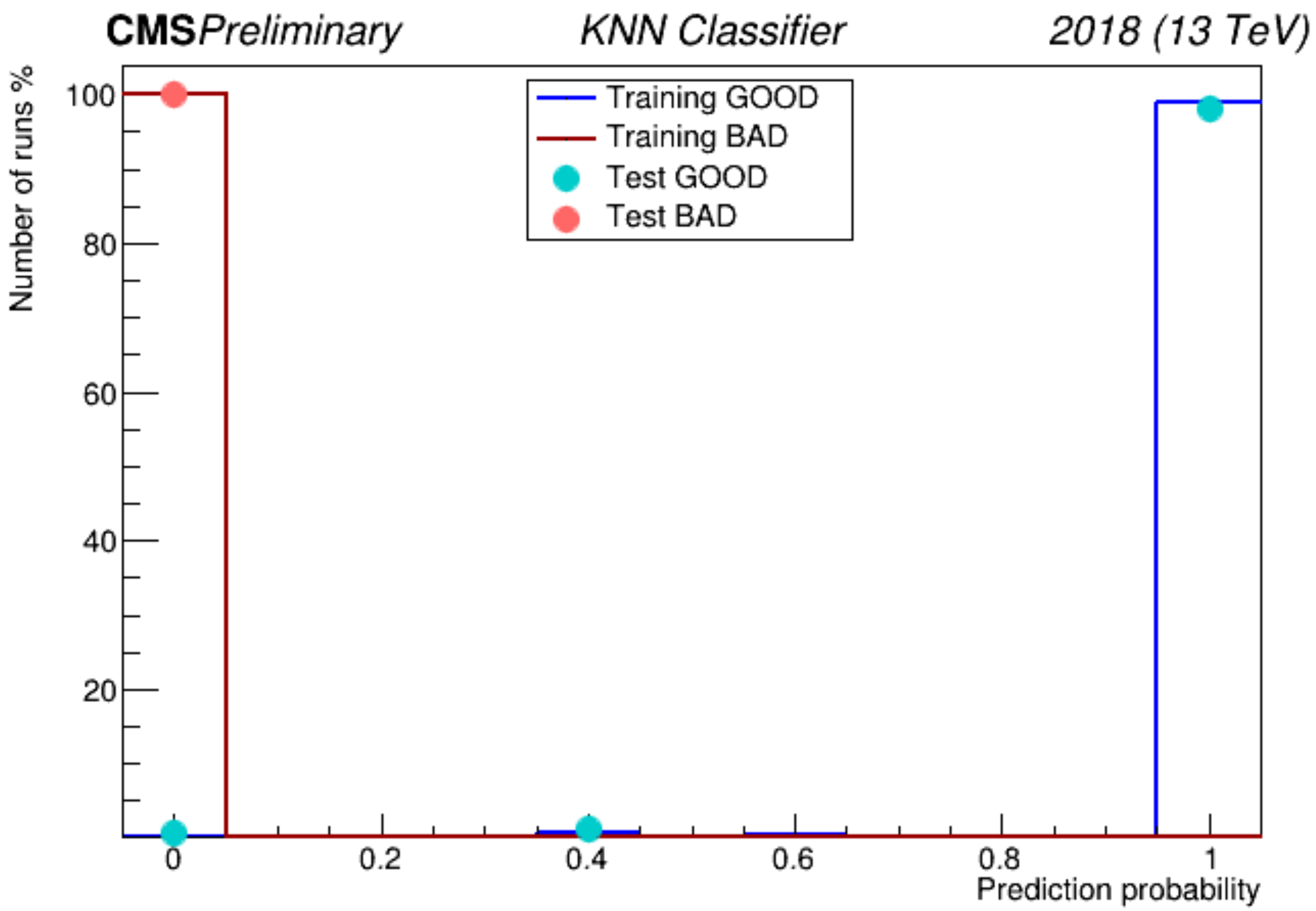


To be updated for Run 3

CMS-DP-2020-019

Supervised Networks for Data Certification





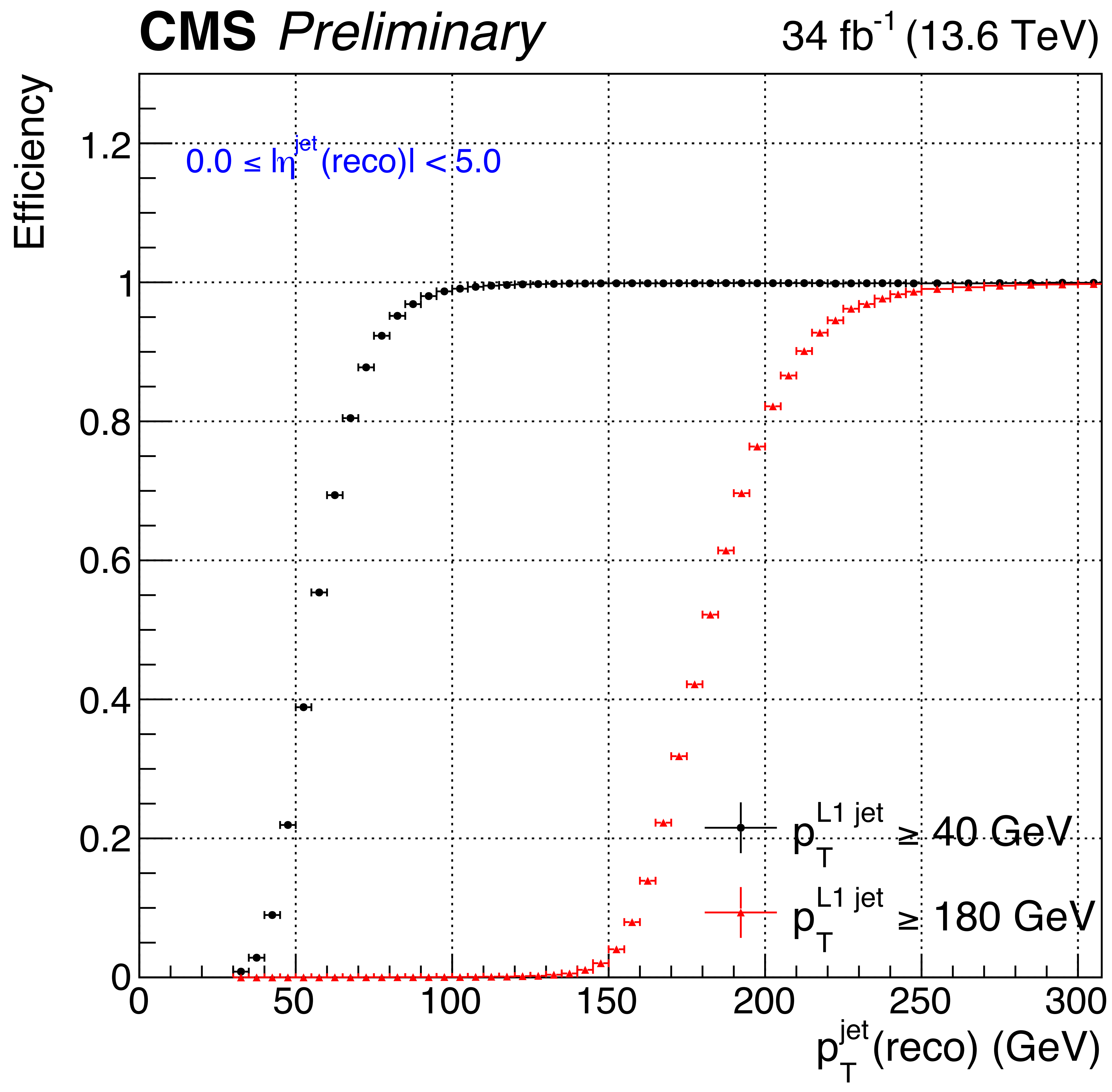
e.g. KNN with 25 input features can achieve ROC AUC of 0.9

Feature reduction: Jet energy fractions among most important for DC

CMS-DP-2023-032

Single Jet L1 trigger

<u>CMS-DP-2023-007</u>



ML for Data certification

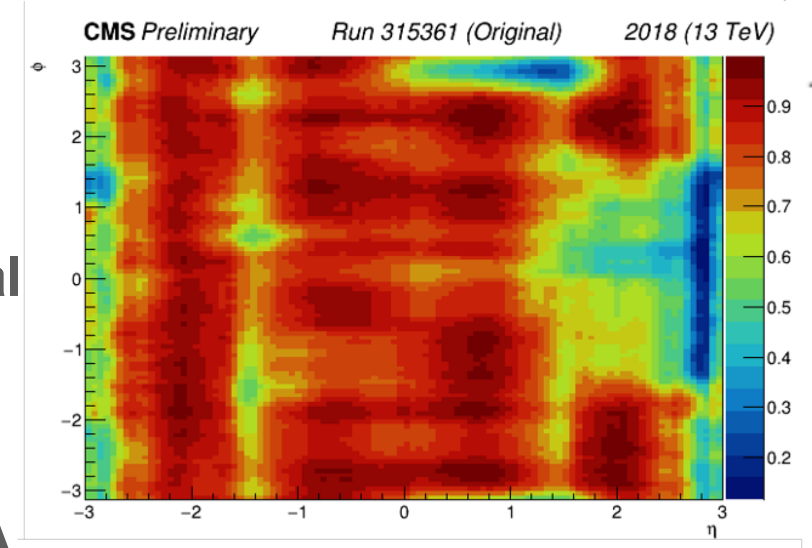
Autoencoders

CMS-DP-2023-032

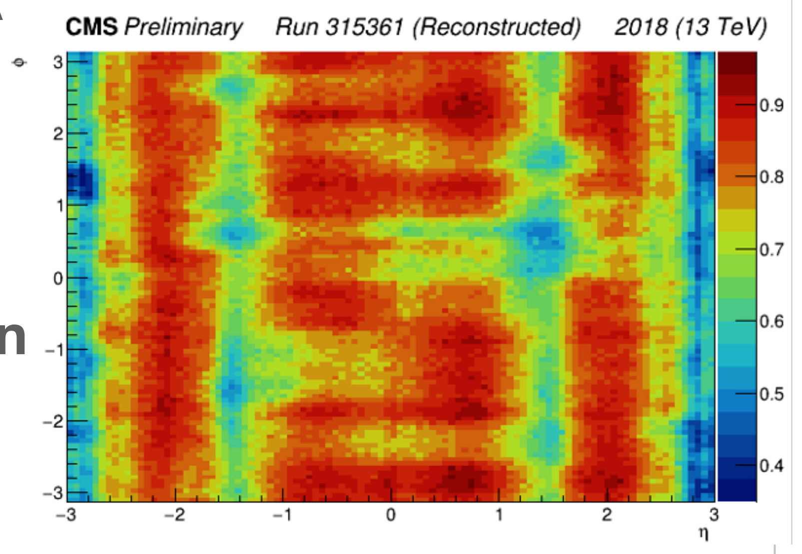
Hadron occupancy

Trained on good runs from Era A

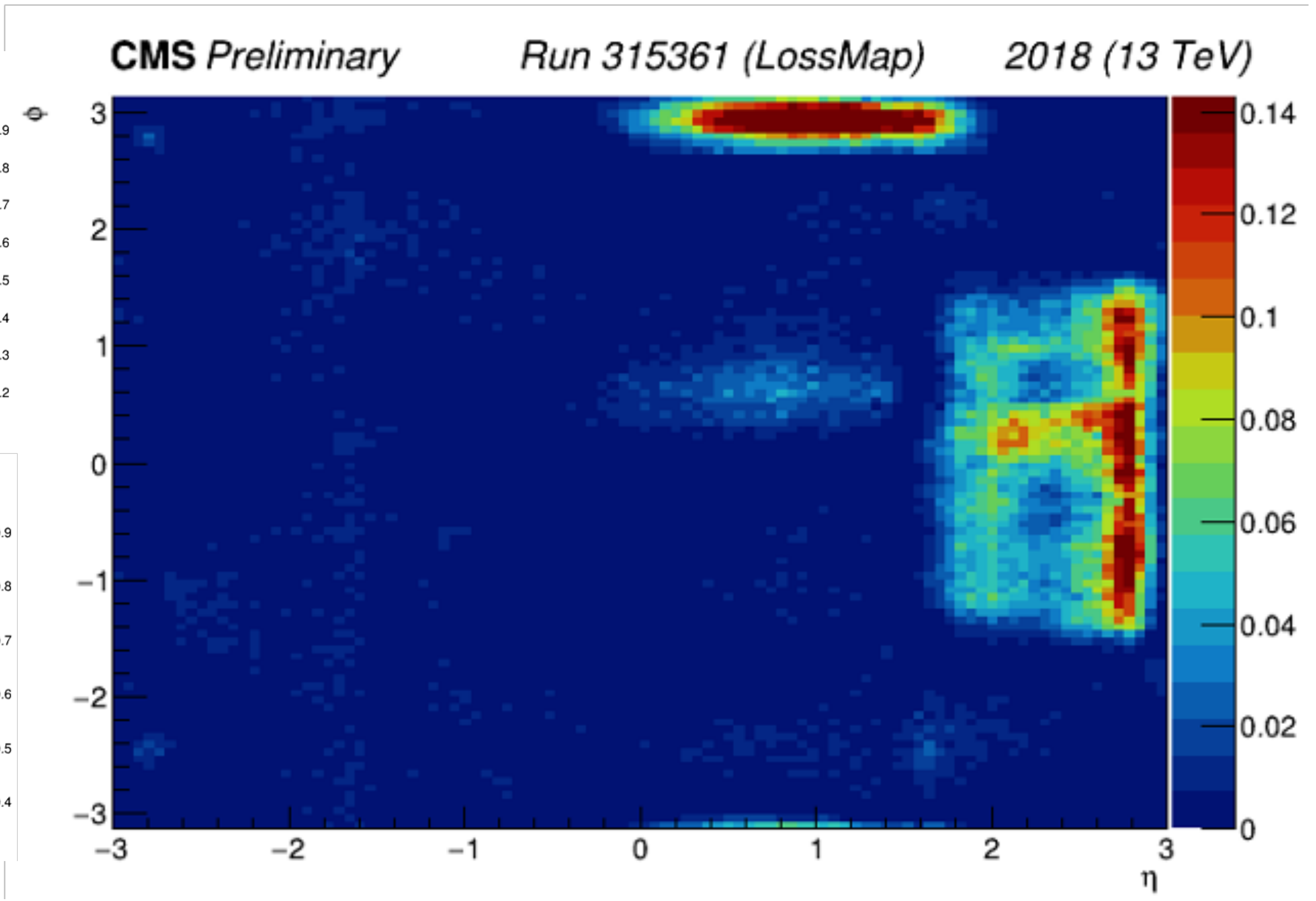
Original



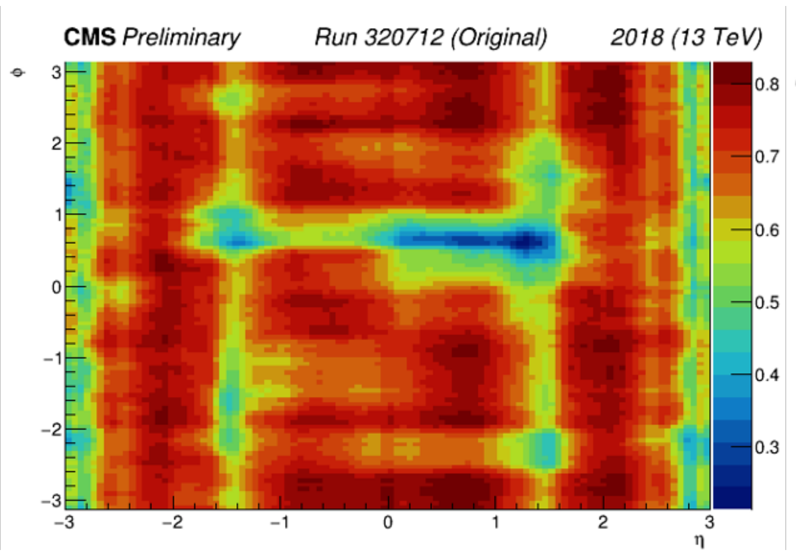
2018 Era A



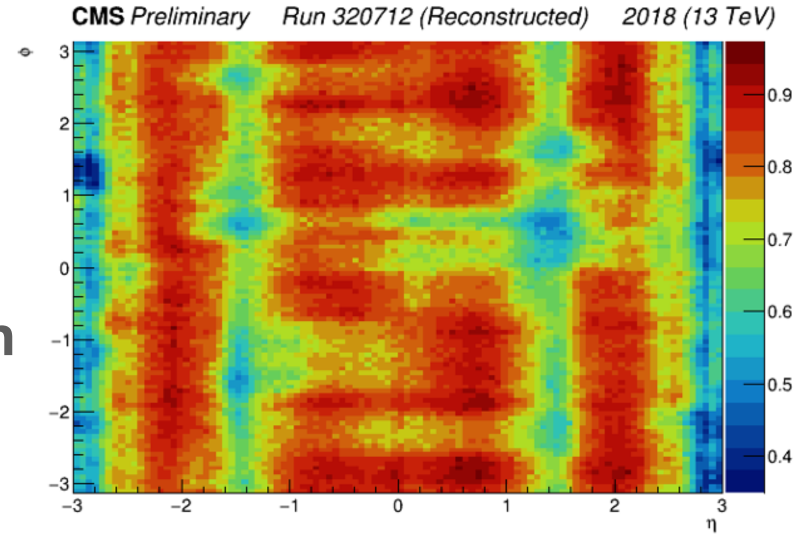
AC reconstruction



Original



2018 Era D



AC reconstruction