Jet reconstruction and calibration: Run 2 and perspective for Run 3

Andrea Malara
on behalf of the CMS Collaboration

2-5 August 2021

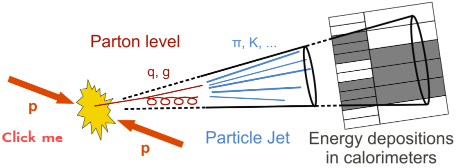
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Bundesministerium für Bildung und Forschung

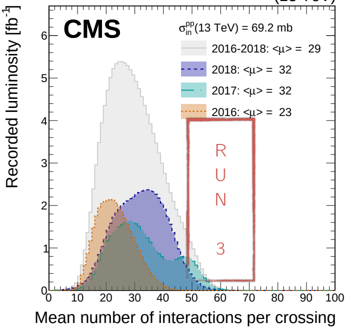
CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE

Motivation

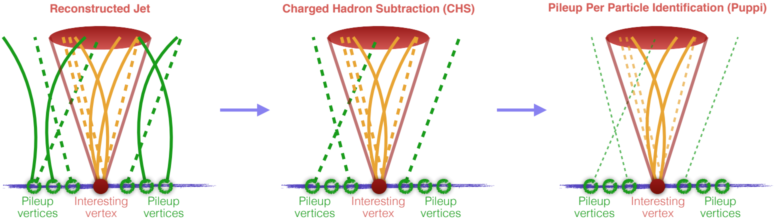
- Abundance of quarks and gluons in pp collisions
- Jets as probe in (almost) all SM and BSM analyses
- Need for precise reconstruction and calibration
- Challenging environment with average pileup of ~ 30 interactions
- Run 3 around the corner, presenting a number of challenges



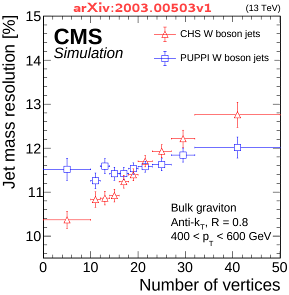
arXiv:2003.00503v1

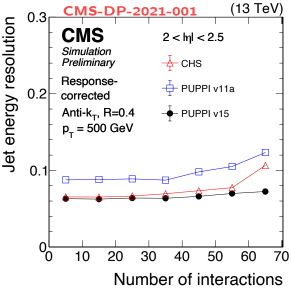


Pileup Mitigation Techniques – CHS vs. PUPPI

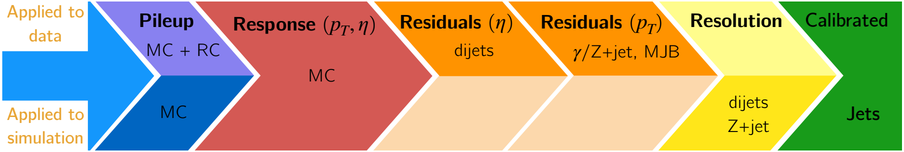


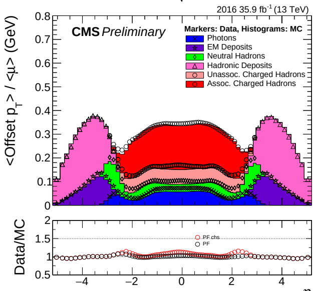
- CHS → charged PF
- Puppi → neutral PF
- Widely used in Run2
- Default in Run3
- Continuous improvement for all jet-related variables





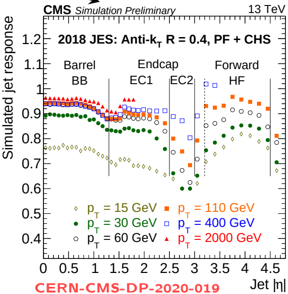
Jet Calibration





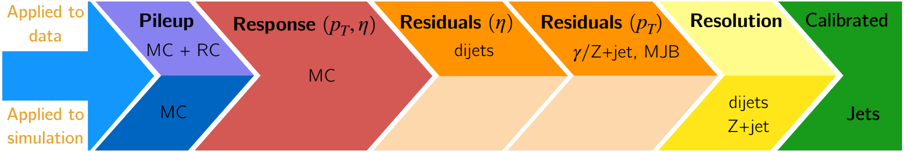
CERN-CMS-DP-2020-019

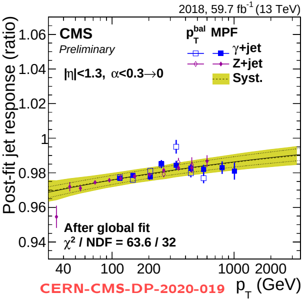
- Simulation-based
- Core of the JEC
- Accounts for detector effects



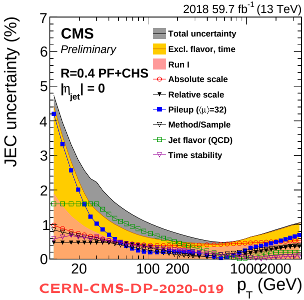
CERN-CMS-DP-2020-019

Jet Calibration





- Residual correction applied to data
- Relative to precisely measured reference objects (μ, e, γ)
- Uncertainty $\sim 1\%$ for jets with $p_T \geq 100$ GeV

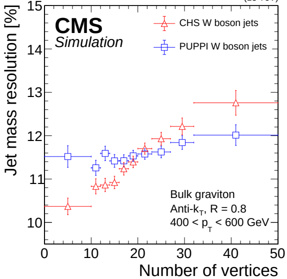


Outlook

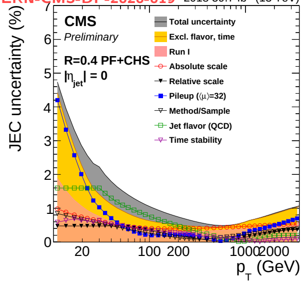
- Run 2 experience fully exploited for accurate jet performance
- Several high-performance methods presented
- Ready to cope with high PU expected for Run 3 (Puppi very promising)
- But it's not the end of the story
  - Run 2 legacy corrections to improve performances (planned <1% JEC unc.)
  - Increasing the granularity of the corrections to tackle detector ageing
  - ML-based approaches will help

More exciting results are yet to come

arXiv:2003.00503v1



CERN-CMS-DP-2020-019

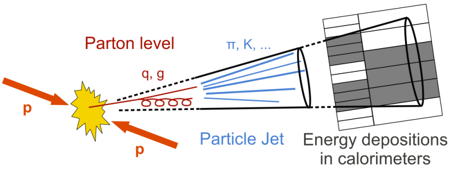


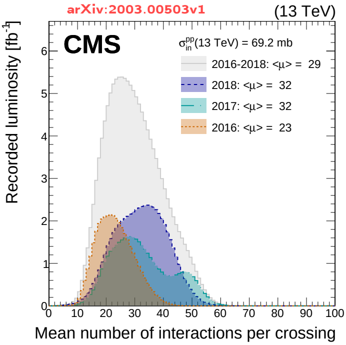
— Backup —

— Full Material —

Jets

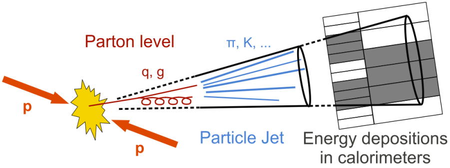
- Abundance of quarks and gluons in pp collisions
- Jets used in (almost) all SM and BSM analyses
- Need for precise reconstruction and calibration
- Challenging environment with average pileup of ~ 30 interactions
- Run 3 around the corner, presenting a number of challenges

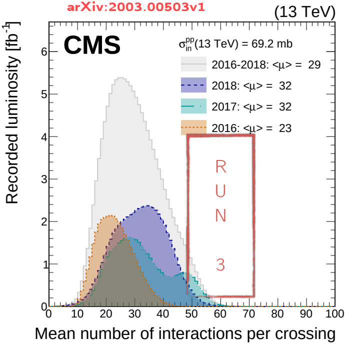




Jets

- Abundance of quarks and gluons in pp collisions
- Jets used in (almost) all SM and BSM analyses
- Need for precise reconstruction and calibration
- Challenging environment with average pileup of ~ 30 interactions
- Run 3 around the corner, presenting a number of challenges





Building Jets

Local reconstruction:
Tracks, ECAL, HCAL

→

- Information from sub-detectors
- Fast in order to cope with PU

Building Jets

Local reconstruction: Tracks, ECAL, HCAL →
- Information from sub-detectors
- Fast in order to cope with PU

↓

Particle flow (PF) →
- Linking of sub-detectors
- Particle identification

Building Jets

Local reconstruction: Tracks, ECAL, HCAL →
- Information from sub-detectors
- Fast in order to cope with PU

↓

Particle flow (PF) →
- Linking of sub-detectors
- Particle identification

↓

Pileup mitigation →
- CHS
- Puppi

Building Jets

Local reconstruction: Tracks, ECAL, HCAL →
- Information from sub-detectors
- Fast in order to cope with PU

↓

Particle flow (PF) →
- Linking of sub-detectors
- Particle identification

↓

Pileup mitigation →
- CHS
- Puppi

↓

Jet clustering →
- Algorithms (AK, CA)
- Cone radii (0.4, 0.8, 1.5)

Building Jets

Local reconstruction: Tracks, ECAL, HCAL →
- Information from sub-detectors
- Fast in order to cope with PU

↓

Particle flow (PF) →
- Linking of sub-detectors
- Particle identification

↓

Pileup mitigation →
- CHS
- Puppi

↓

Jet clustering →
- Algorithms (AK, CA)
- Cone radii (0.4, 0.8, 1.5)

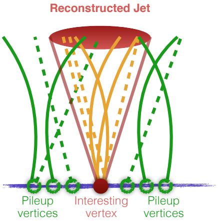
↓

Jet calibration →
- Jet energy scale & resolution
- Jet mass resolution

Pileup Mitigation Techniques

Pileup

- Multiple interactions during a bunch crossing
- Additional particles deteriorate measurements
- Major challenge in LHC physics
- Several approaches to cope with up to 200 interactions



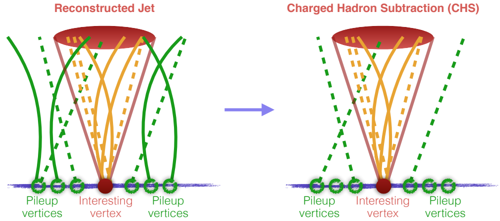
Pileup Mitigation Techniques

Pileup

- Multiple interactions during a bunch crossing
- Additional particles deteriorate measurements
- Major challenge in LHC physics
- Several approaches to cope with up to 200 interactions

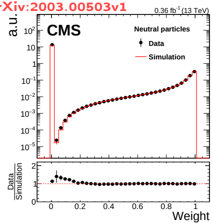
Charged Hadron Subtraction (CHS)

- Tracker information to remove charged particles associated to PU
- Applicable for $|\eta| < 2.4$



Pileup Mitigation Techniques

arXiv:2003.00503v1



Pileup

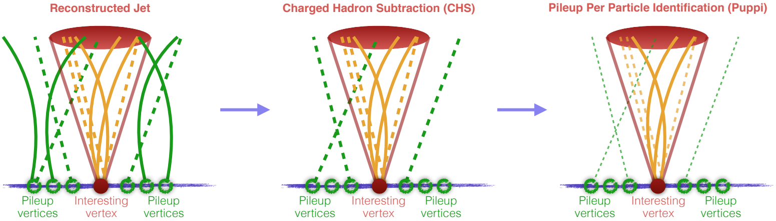
- Multiple interactions during a bunch crossing
- Additional particles deteriorate measurements
- Major challenge in LHC physics
- Several approaches to cope with up to 200 interactions

Charged Hadron Subtraction (CHS)

- Tracker information to remove charged particles associated to PU
- Applicable for $|\eta| < 2.4$

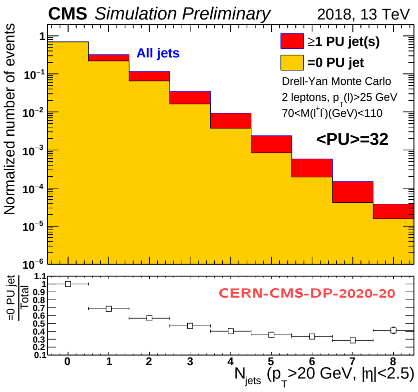
Pileup Per Particle Identification (Puppi)

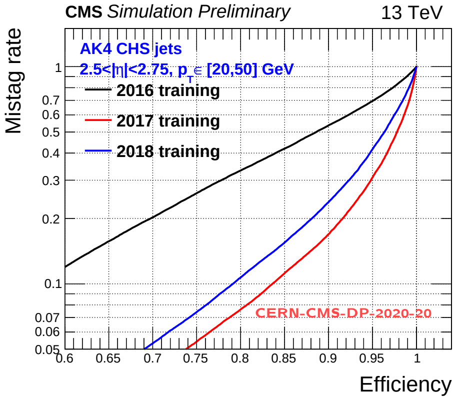
- Per-particle weight
- Scale 4-momentum before clustering
- Charged particles similar to CHS



Pileup Mitigation Techniques – PU jet ID

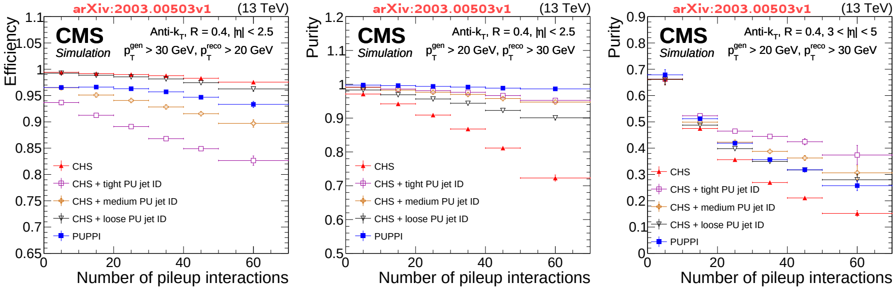
- BDT-based discriminator to identify low-p_T jets coming from PU
- High rejection power and low mistag rate at 95% efficiency
- Large improvement from Phase-1 upgrade of pixel detector after 2016: extended coverage from $|\eta| < 2.5$ to 2.7





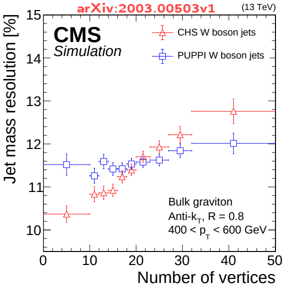
Pileup Mitigation Techniques – CHS vs. PUPPI

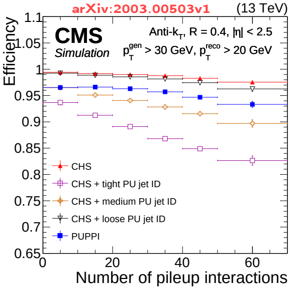
- Puppi: Good balance between purity and efficiency for $|\eta| < 2.5$
- CHS: Improves purity at cost of efficiency only in combination with PU jet ID

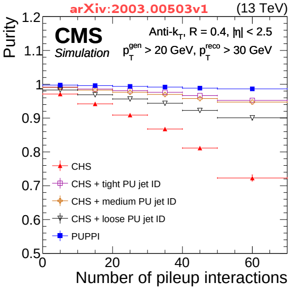


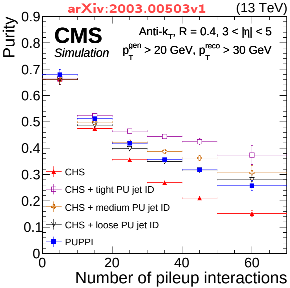
Pileup Mitigation Techniques – CHS vs. PUPPI

- Puppi: Good balance between purity and efficiency for $|\eta| < 2.5$
- CHS: Improves purity at cost of efficiency only in combination with PU jet ID
- Puppi: Jet variables more stable against PU (e.g. mass resolution)



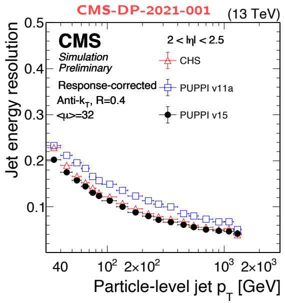


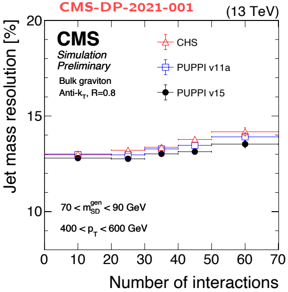


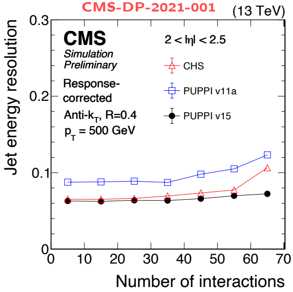


Pileup Mitigation Techniques – Puppi for Run 3

- Widely used in Run 2 (Puppi v11a)
- Refined requirements for charged particles (Puppi v15)
- Targeting better jet energy resolution at high-p_T
- Improved all jet-related variables
- Used for Run 2 Legacy reconstruction
- Default in Run 3



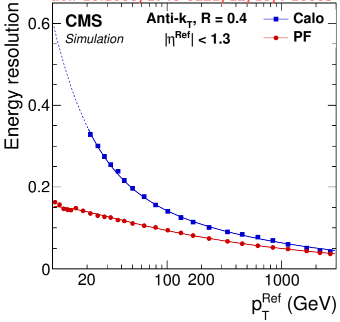


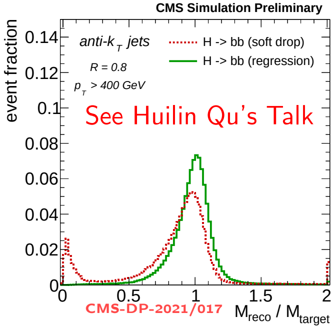


Jet Reconstruction

- Clustering of PF particles
  - successfully used since Run1
  - better performance than calorimeter-only-based reconstruction
- Anti-k_T as default algorithm
  - infrared and collinear safe
  - alternative algorithms: HOTVR, XCone

doi: 10.1088/1748-0221/12/10/P10003





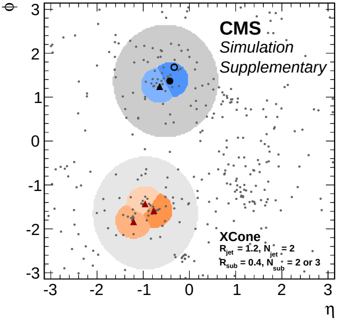
CMS-DP-2021/017

- Typical jet collections in CMS:
  - small radius: R=0.4 (AK4)
  - large radius: R=0.8 (AK8)
- Substructure for AK8:
  - key role in jet tagging
  - softdrop algorithm for mass regression

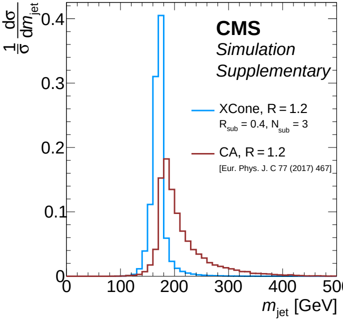
Jet Reconstruction – XCone

- Event signature defines clustering
- Return exactly N jets
- Examples from top-mass measurement
- Large improvement for the jet mass resolution

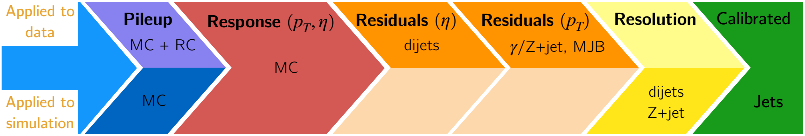
Phys. Rev. Lett. 124 (2020) 202001



Phys. Rev. Lett. 124 (2020) 202001

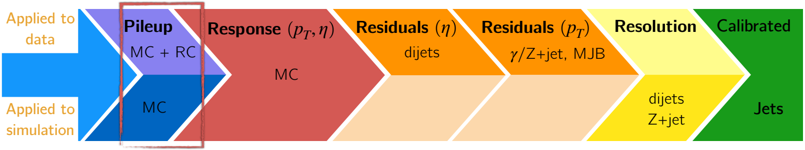


Jet Calibration



Strategy: Factorized approach

Jet Calibration

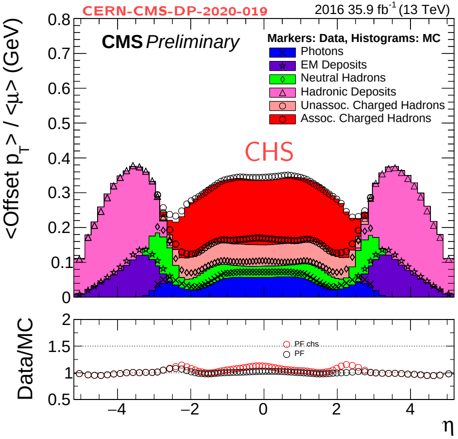


Strategy: Factorized approach

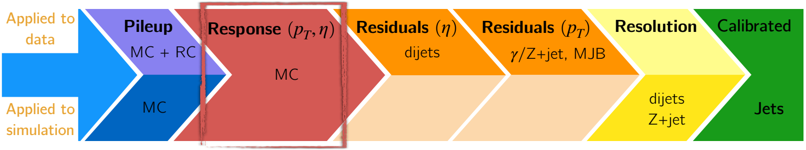
MC truth correction:

PU subtraction

- Average difference (offset) in p_T between matched jets
- Simulation-based corrections (multijet sample w/ and w/o PU) applied to data and simulation
- Residual corrections derived with Random Cone algorithm applied to data
- Monitored for each type of PF candidates



Jet Calibration



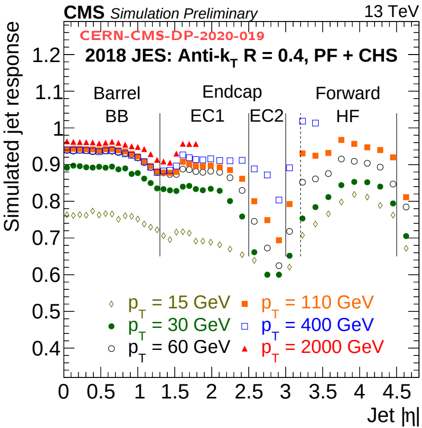
Strategy: Factorized approach

MC truth correction:

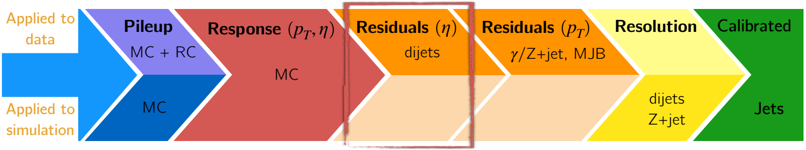
PU subtraction

Jet response calibration

- Core of the JEC
- Measured in simulation and applied to data as well
- Accounts for detector effects
- Change in performance at high $|\eta|$ and low p_T due to detector acceptance



Jet Calibration



Strategy: Factorized approach

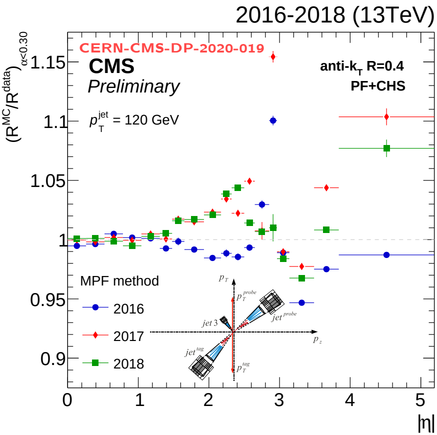
MC truth correction:

PU subtraction

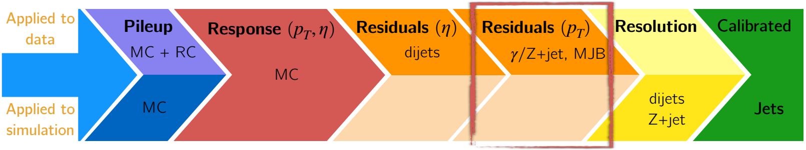
Jet response calibration

Residual corrections

- Small residual correction of jet response applied to data
- Address different response in each sub-detector (η dep.)
- Sizeable corrections in detector transition regions



Jet Calibration



Strategy: Factorized approach

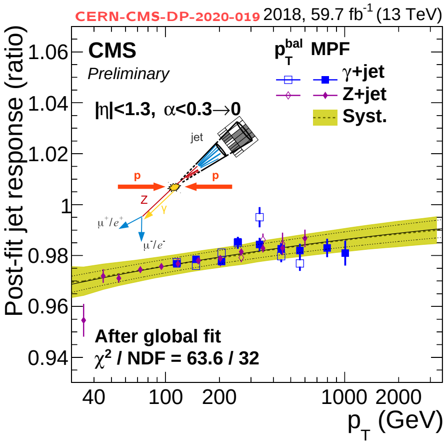
MC truth correction:

PU subtraction

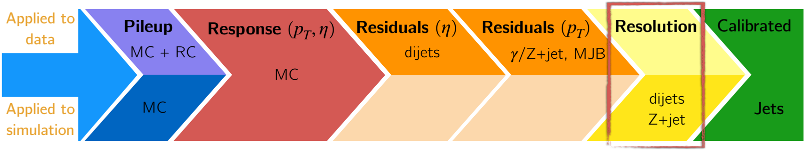
Jet response calibration

Residual corrections

- Additional p_T dep. corrections accounting for abs. scale in barrel
- Determined relative to precisely measured reference objects (μ, e, γ)
- Combined in a global fit (reference object scales as nuisance parameters)



Jet Calibration



Strategy: Factorized approach

MC truth correction:

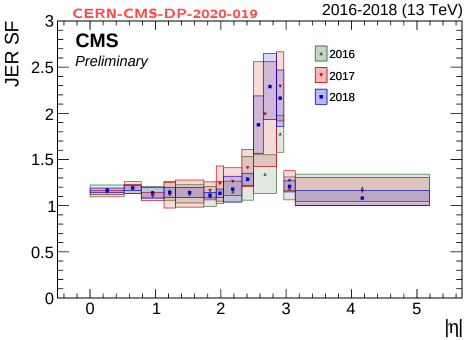
PU subtraction

Jet response calibration

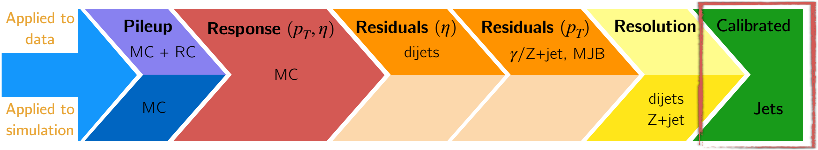
Residual corrections

Jet energy resolution smearing

- Scale factors (SFs) applied to simulation to match jet resolution in data
- p_T and η dependent SFs aim for coverage of full phase space



Jet Calibration



Strategy: Factorized approach

MC truth correction:

PU subtraction

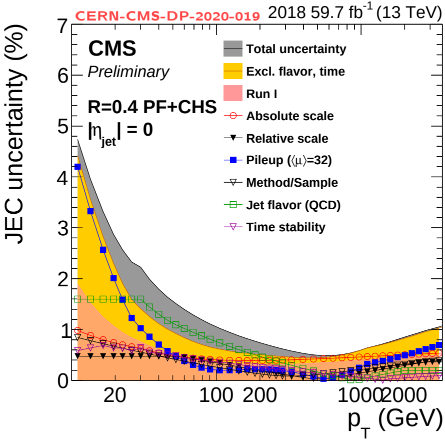
Jet response calibration

Residual corrections

Jet energy resolution smearing

Jet Energy Scale Uncertainty

- Uncertainty $\sim 1\%$ for jets with $p_T \geq 100$ GeV
- Large contribution from PU corrections

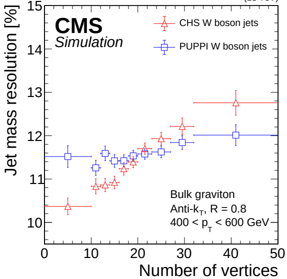


Outlook

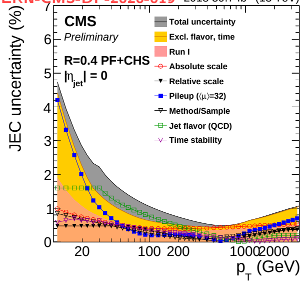
- Run 2 experience fully exploited for accurate jet performance
- Several high-performance methods presented
- Puppi very promising to cope with high PU expected for Run 3
- But it's not the end of the story
  - Run 2 legacy corrections to improve performances (planned <1% JEC unc.)
  - Increasing granularity of corrections to tackle detector ageing
  - ML-based approaches will help

More exciting results are yet to come

arXiv:2003.00503v1



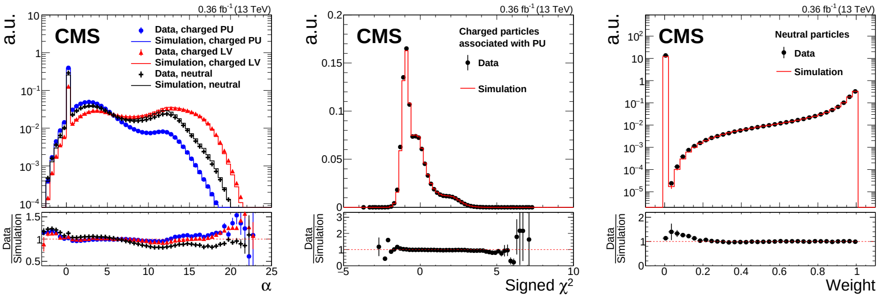
CERN-CMS-DP-2020-019



—Additional Material—

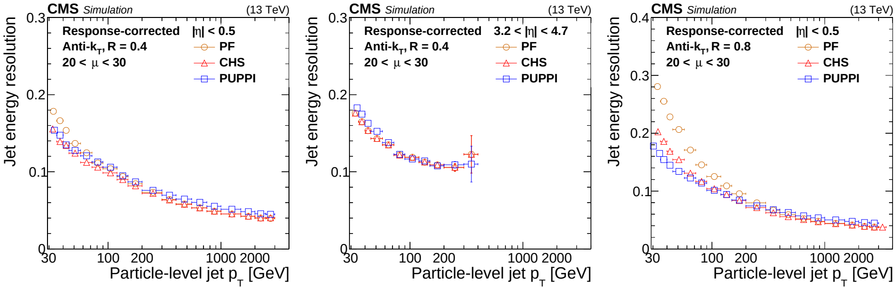
Puppi – Weights

arXiv:2003.00503v1



Puppi – Jet Resolution vs p_T

arXiv:2003.00503v1

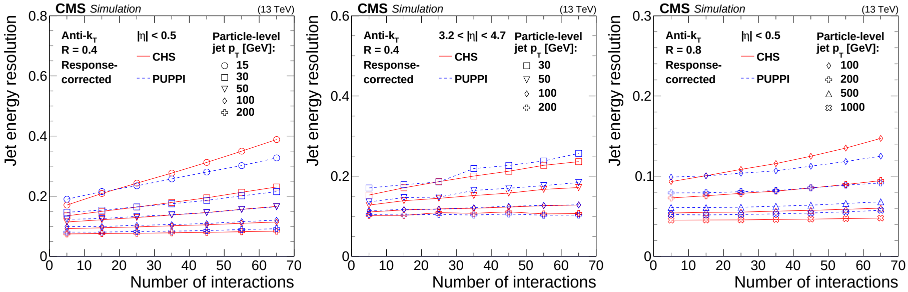


Significant gain with PUPPI particularly at low-p_T jets

Updated PUPPI tune in place to improve jet resolution at high-p_T

Puppi – Jet Resolution vs PU

arXiv:2003.00503v1

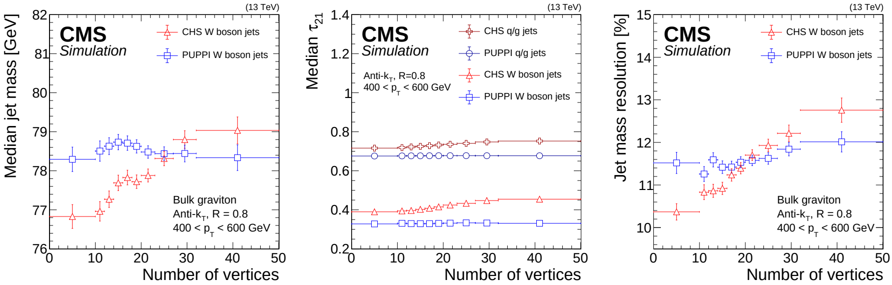


Significant gain with PUPPI particularly at low-p_T jets

Updated PUPPI tune in place to improve jet resolution at high-p_T

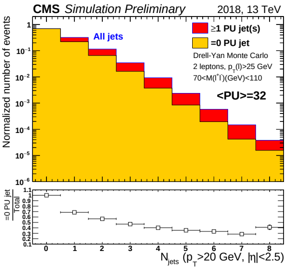
Puppi – Jet Variables

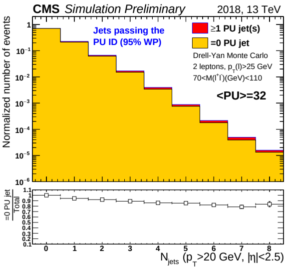
arXiv:2003.00503v1

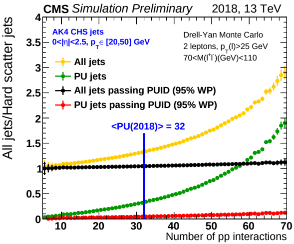


PU jet ID

CERN-CMS-DP-2020-20



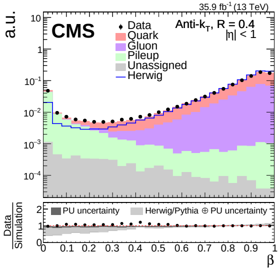


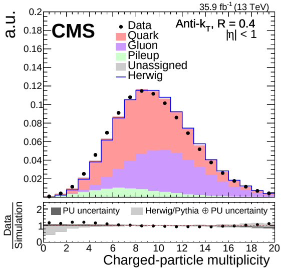


PU jet ID – Input Variables

arXiv:2003.00503v1

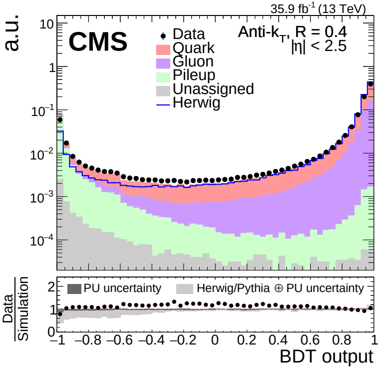
Input variable	Definition
β	Fraction of p_{T} of charged particles associated with the LV, defined as $\sum_{i\in \mathrm{LV}} p_{\mathrm{T},i} / \sum_i p_{\mathrm{T},i}$ where i iterates over all charged PF particles in the jet
N_{vertices}	Number of vertices in the event
$\langle \Delta R^2 \rangle$	Square distance from the jet axis scaled by p_{T}^2 average of jet constituents: $\sum_i \Delta R^2 p_{\mathrm{T},i}^2 / \sum_i p_{\mathrm{T},i}^2$
$f_{\mathrm{ring}X}$, $X = 1, 2, 3,$ and 4	Fraction of p_{T} of the constituents ($\sum p_{\mathrm{T},i}/p_{\mathrm{T}}^{\mathrm{jet}}$) in the region $R_i < \Delta R < R_{i+1}$ around the jet axis, where $R_i = 0, 0.1, 0.2,$ and 0.3 for $X = 1, 2, 3,$ and 4
$p_{\mathrm{T}}^{\mathrm{lead}} / p_{\mathrm{T}}^{\mathrm{jet}}$	p_{T} fraction carried by the leading PF candidate
$p_{\mathrm{T}}^{\mathrm{l.\,ch.}} / p_{\mathrm{T}}^{\mathrm{jet}}$	p_{T} fraction carried by the leading charged PF candidate
$\lvert \vec{m} \rvert$	Pull magnitude, defined as $\lvert(\sum_i p_{\mathrm{T}}^i \lvert r_i \rvert \vec{r}_i)\rvert / p_{\mathrm{T}}^{\mathrm{jet}}$ where $\vec{r}_i$ is the direction of the particle i from the direction of the jet
N_{total}	Number of PF candidates
N_{charged}	Number of charged PF candidates
σ_1	Major axis of the jet ellipsoid in the η-ϕ space
σ_2	Minor axis of the jet ellipsoid in the η-ϕ space
$p_{\mathrm{T}}^{\mathrm{D}}$	Jet fragmentation distribution, defined as $\sqrt{\sum_i p_{\mathrm{T},i}^2} / \sum_i p_{\mathrm{T},i}$

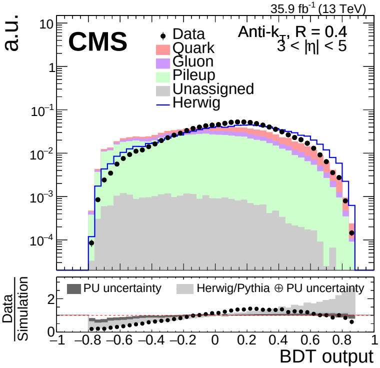




PU jet ID – BDT Output

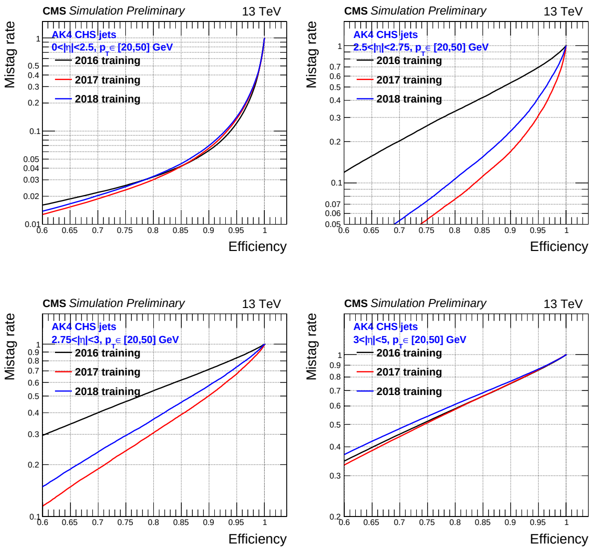
arXiv:2003.00503v1





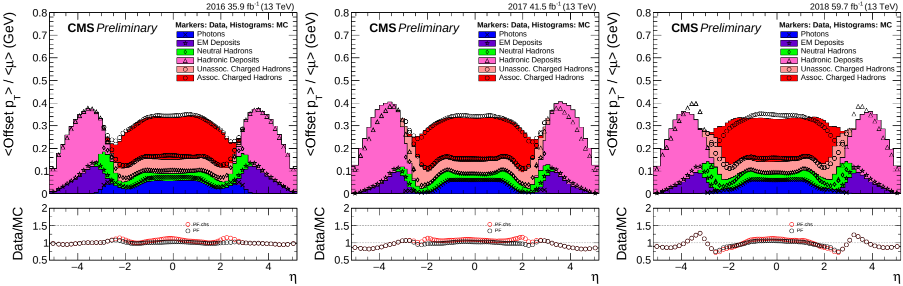
PU jet ID – ROC

CERN-CMS-DP-2020-20



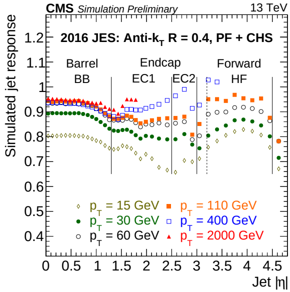
PU Offset

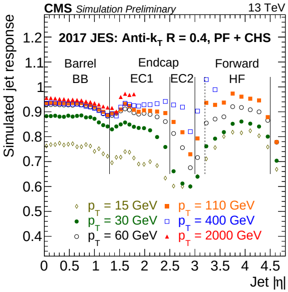
CERN-CMS-DP-2020-019

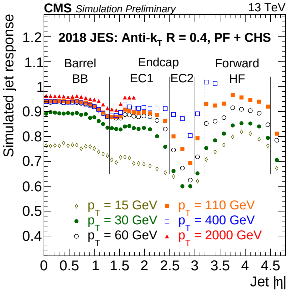


Jet Response

CERN-CMS-DP-2020-019

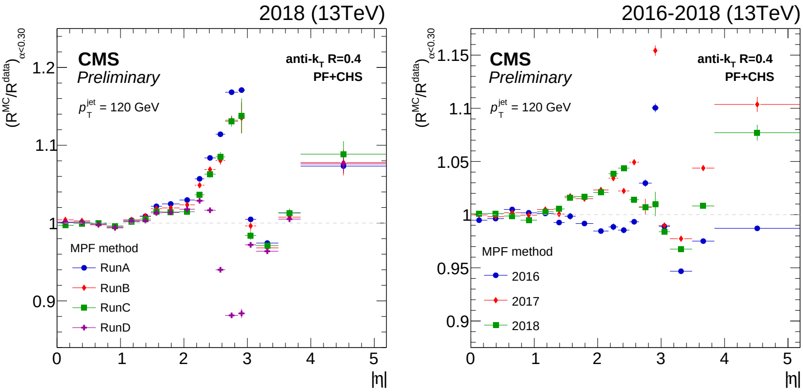






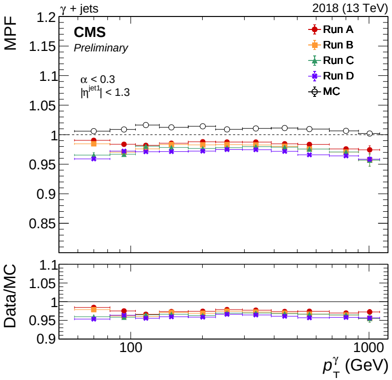
Residuals – η

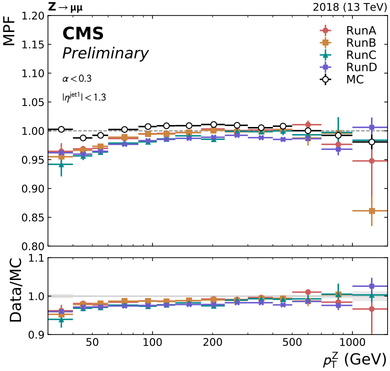
CERN-CMS-DP-2020-019



Residuals – p_T

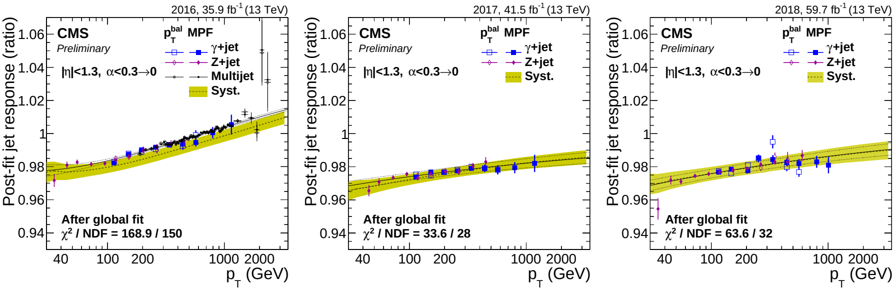
CERN-CMS-DP-2020-019





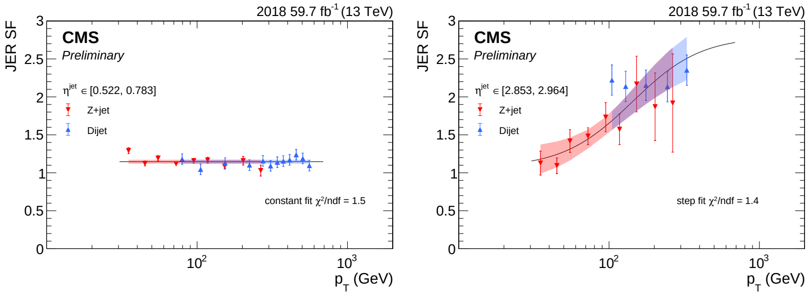
Residuals – Global fit

CERN-CMS-DP-2020-019



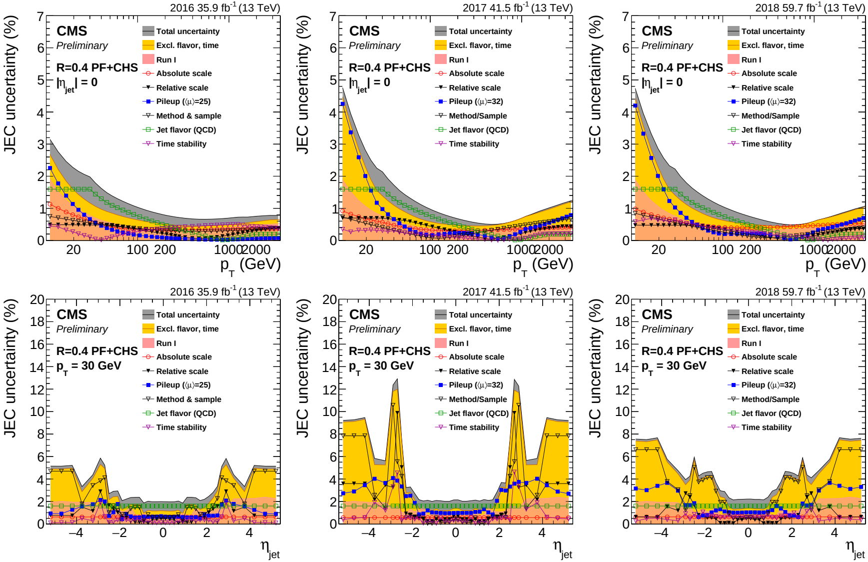
JER SF

CERN-CMS-DP-2020-019



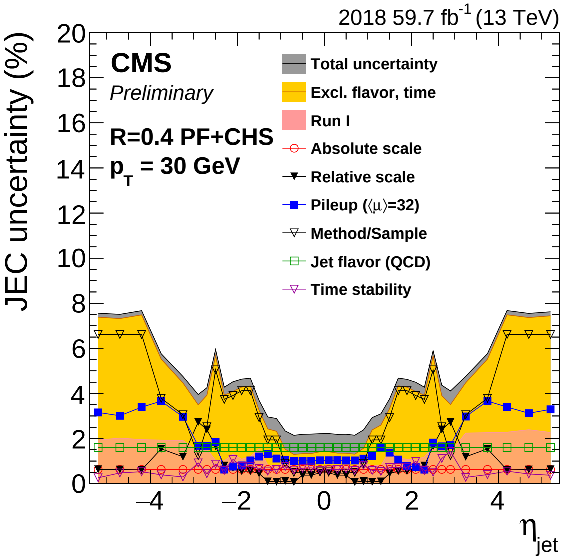
JEC Uncertainty

CERN-CMS-DP-2020-019



JEC Uncertainty

CERN-CMS-DP-2020-019



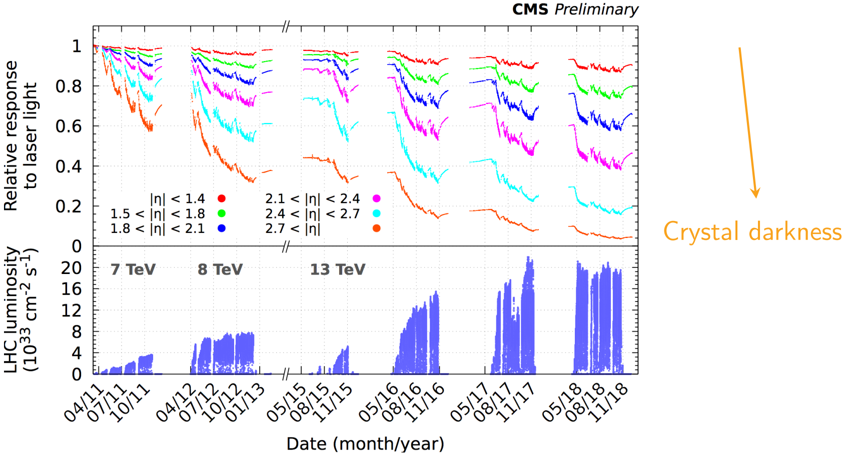
- Issues impacting uncertainty:
  - PU
  - Detector ageing (ECAL transparency loss)
  - Malfunctions (TKR damage)

Run 2 ($\sim 2-8\%$)

Run 1 ($\sim 1-2\%$)

ECAL

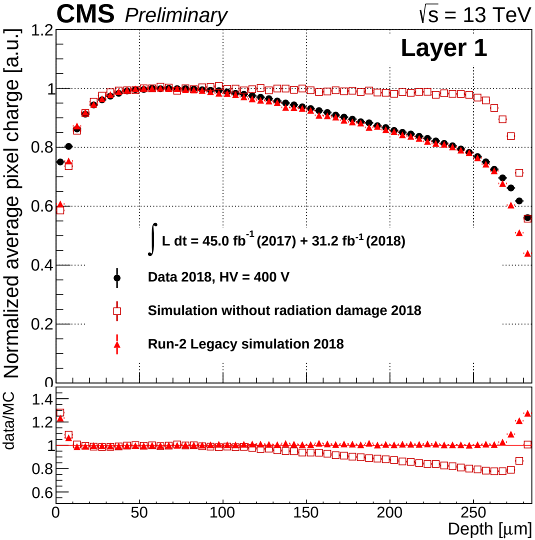
CMS-CR-2019-069



Evolution of crystal transparency in different regions of the detector during LHC Run 1 and Run 2

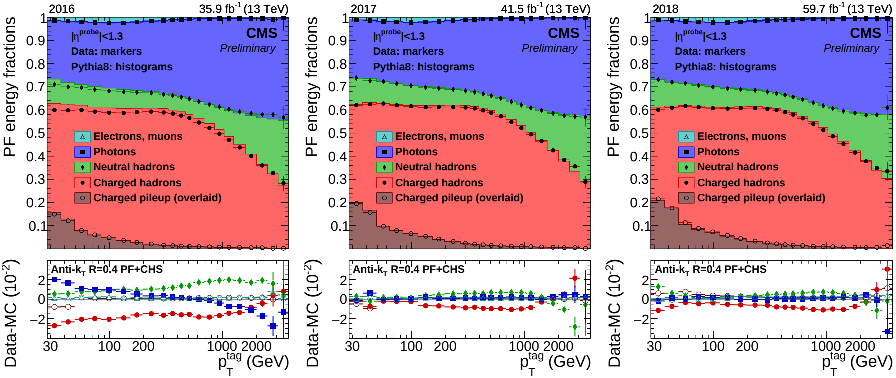
PIX/TRK radiation damage

CERN-CMS-DP-2020-026



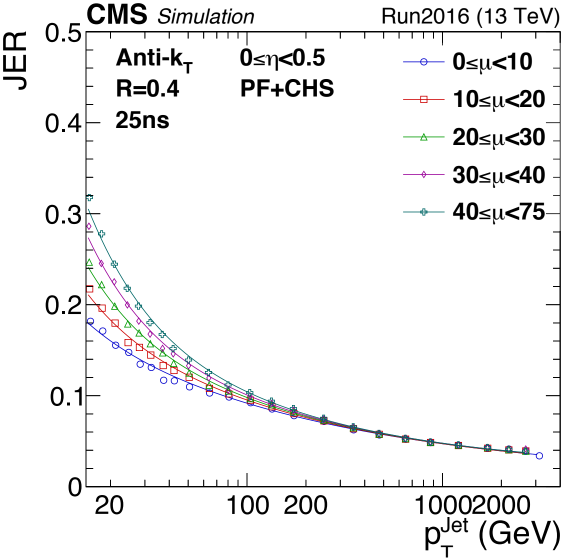
PF composition

CERN-CMS-DP-2020-019



JER

CERN-CMS-DP-2016-020



Run 1

Run 2

Run 3

ML for Hbb regression – Resolution

CMS-PAS-HIG-18-027