Pileup mitigation in jets with CMS

Ksenia de Leo, Anna Benecke, Andreas Hinzmann, Roman Kogler

University of Hamburg

UHH group meeting

January 15, 2020

CLUSTER OF EXCELLENCE

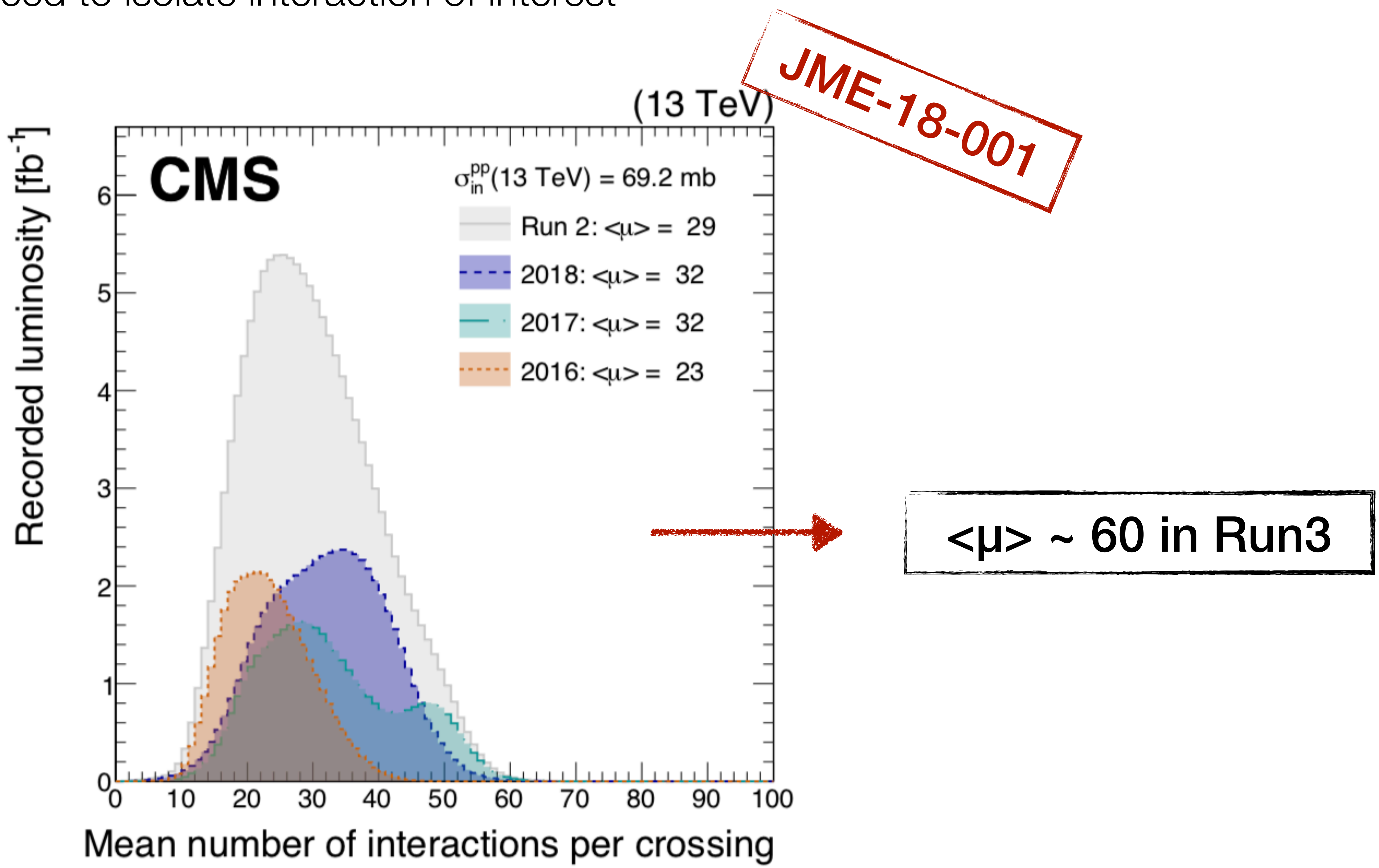
QUANTUM UNIVERSE

Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Pileup in Run2

- In each bunch crossing: **main interaction** + additional collisions (**pileup**)
- Need to isolate interaction of interest



Pileup Mitigation Techniques

- Charged Hadron Subtraction **CHS**
- Pileup Per Particle Identification **PUPPI**
- Pileup Jet ID: reject PU jets (applied on CHS)

CHS

- **Charged particles** associated to PU vertex are removed from jets
- All **neutral particles** are kept
- Additional correction on jets to treat neutral PU

PUPPI Algorithm

- Each particle is assigned a weight **$0<w<1$**:
  - 0 it is most probably originating from PU
  - 1 it is most probably originating from LV
- Weight applied to 4-momentum of the particle
- **Charged particles** are given a weight of **0 or 1** based on the track-vertex association (fromPV and d_Z variable)
- For **neutral particles** the weight is calculated based on surrounding particles and charged PU distribution in the event

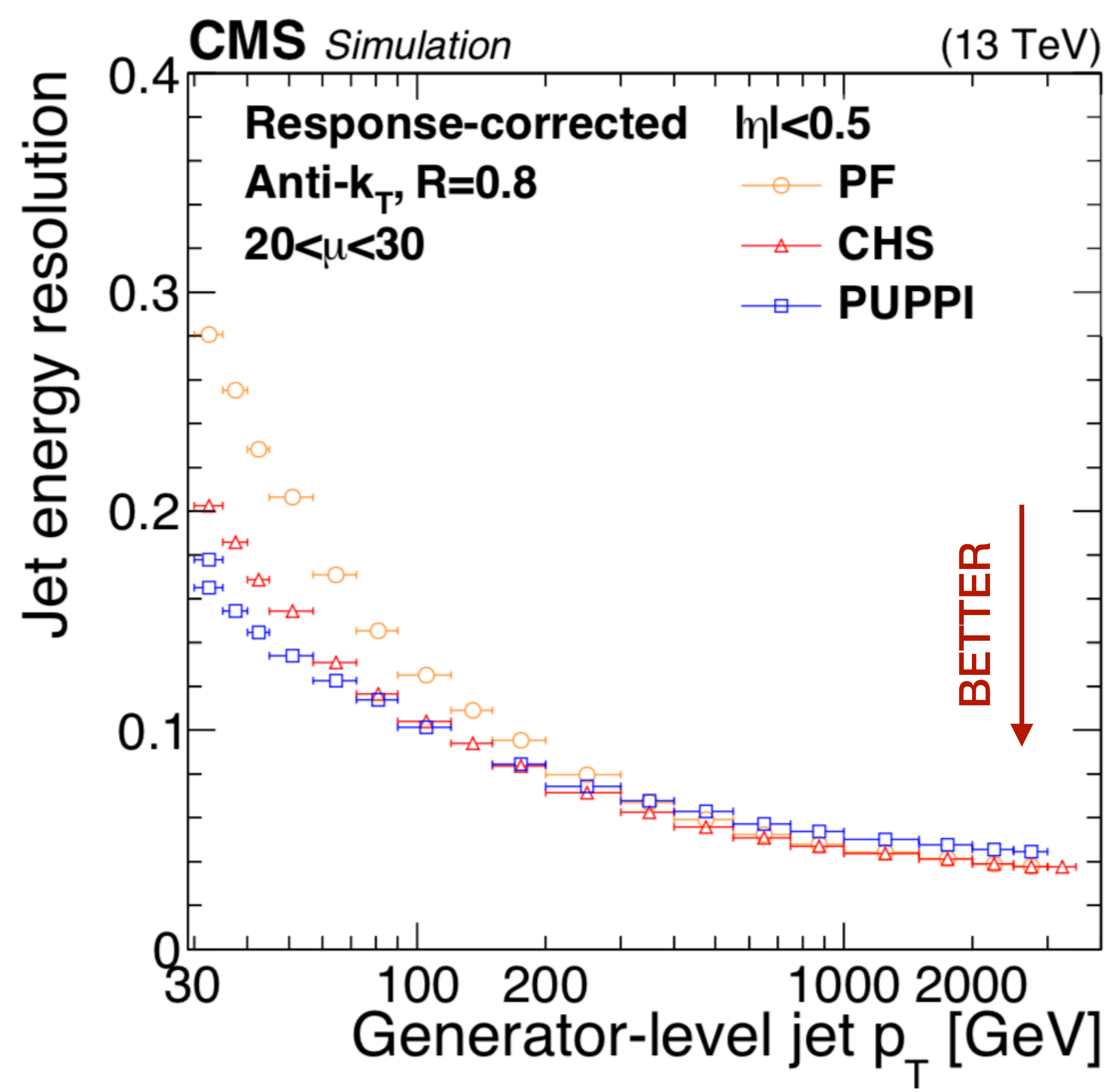
PUPPI Algorithm

- PUPPI shows very good performance **(JME-18-001)**: jets, MET, lepton isolation, jet substructure…

→ **goal**: PUPPI as default algorithm in Run3

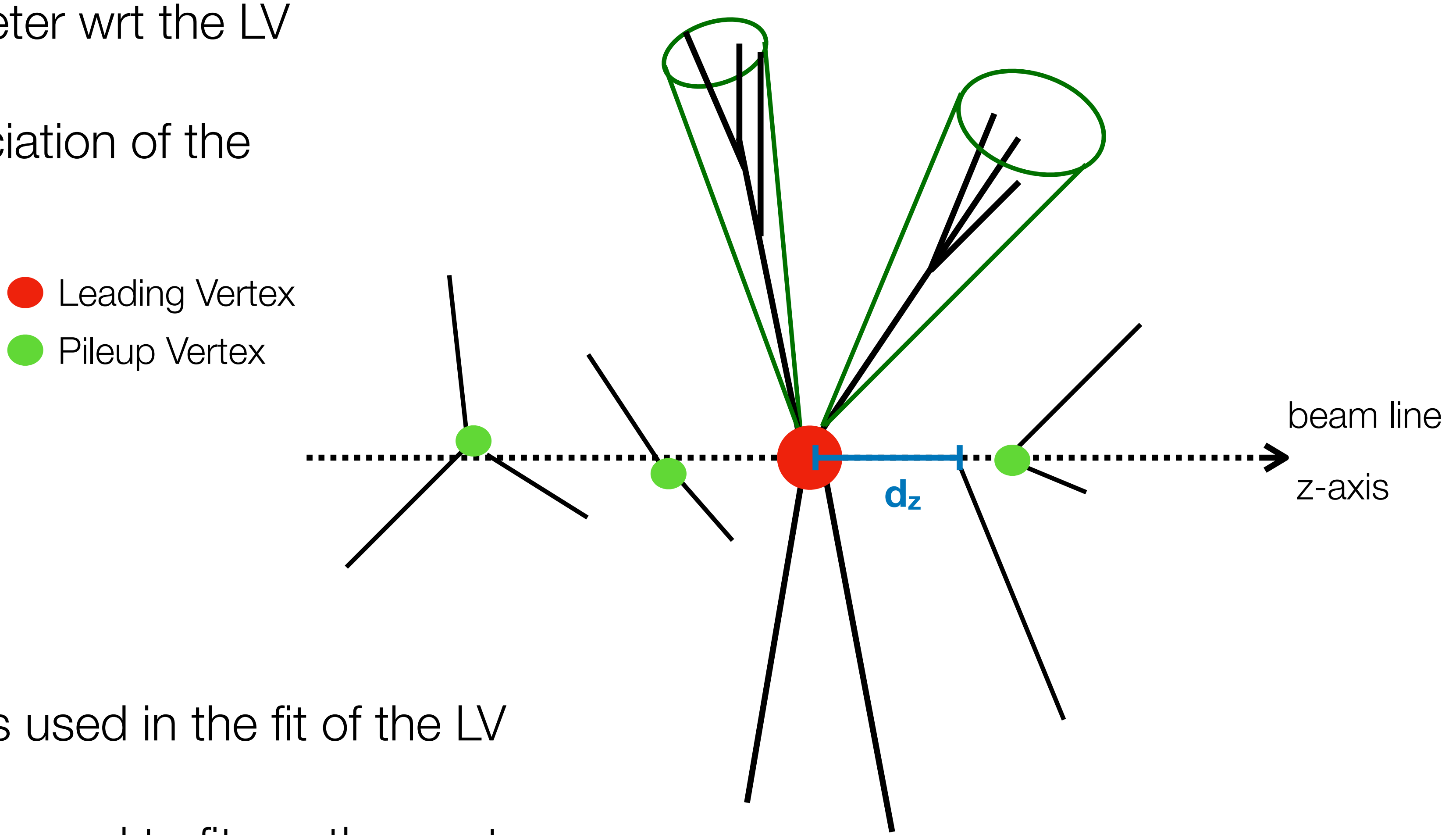
- Improvements: JER at high p_T

At **low p_T** PUPPI shows better performance than CHS, but worse performance at **high p_T**



fromPV & d_Z

- Some charged particles are rejected because of the **fromPV** and **d_Z** variables
- **d_Z** is the transverse impact parameter wrt the LV
- **fromPV** tells how tight is the association of the particle to the LV



- Four values between 3 and 0:

vertex association
  - **fromPV=3** the track is used in the fit of the LV
  - **fromPV=0** the track is used to fit another vertex

no vertex association
  - **fromPV=2** the track is not used in the fit of any vertex and is close in z to the LV
  - **fromPV=1** the track is close in z to another vertex

fromPV & d_Z in CHS and PUPPI

- In CHS the charged particles with **fromPV>0** are kept
- In PUPPI:
  - charged particles from a PU vertex are assigned a weight w=0 and are removed ▹ **fromPV>0** as in CHS
  - charged particles from the LV are assigned a weight w=1 ▹ **fromPV=3**
  - no vertex association, w=1 if **d_Z<0.3** cm, w=0 otherwise ▹ **fromPV=1, 2** & **d_Z<0.3** cm

PROBLEM ➜ charged particles at high p_T are ***not well assigned to the vertex***

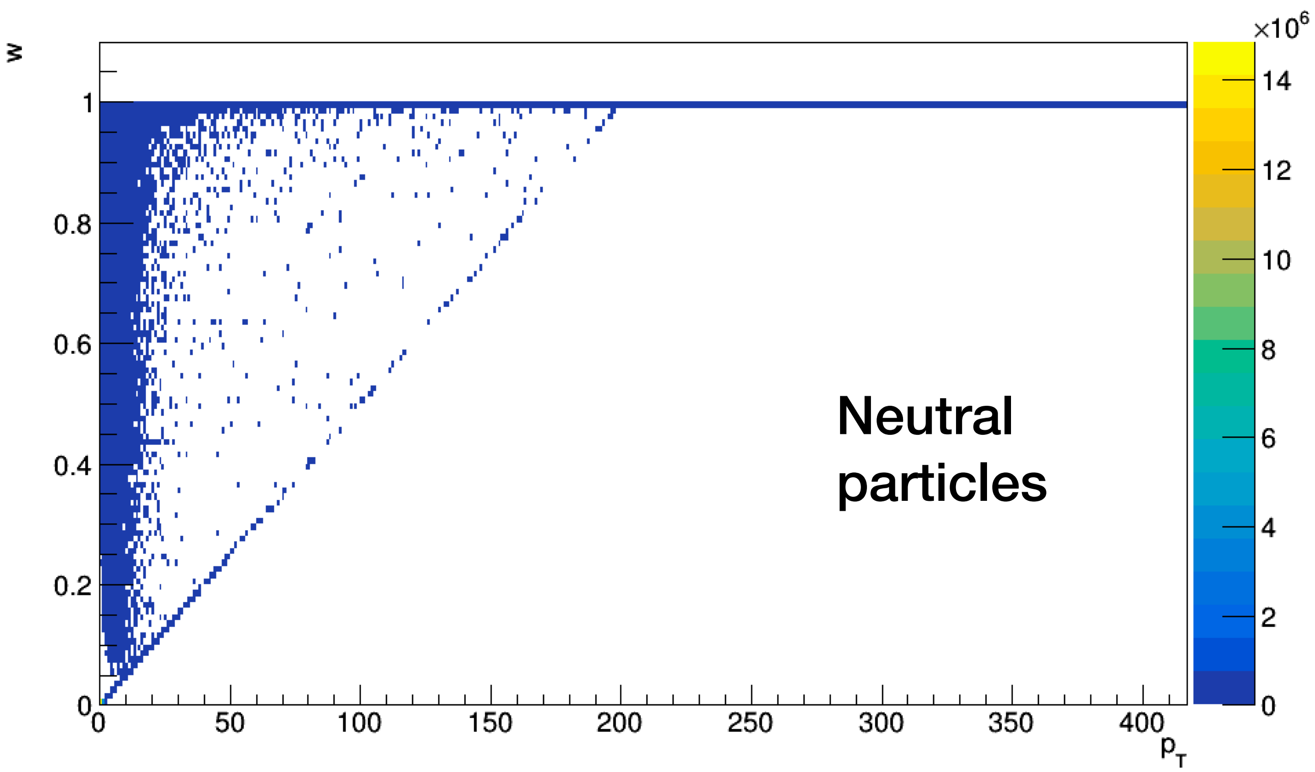
- In the following we will work only with charged particles with **fromPV=1 & 2**

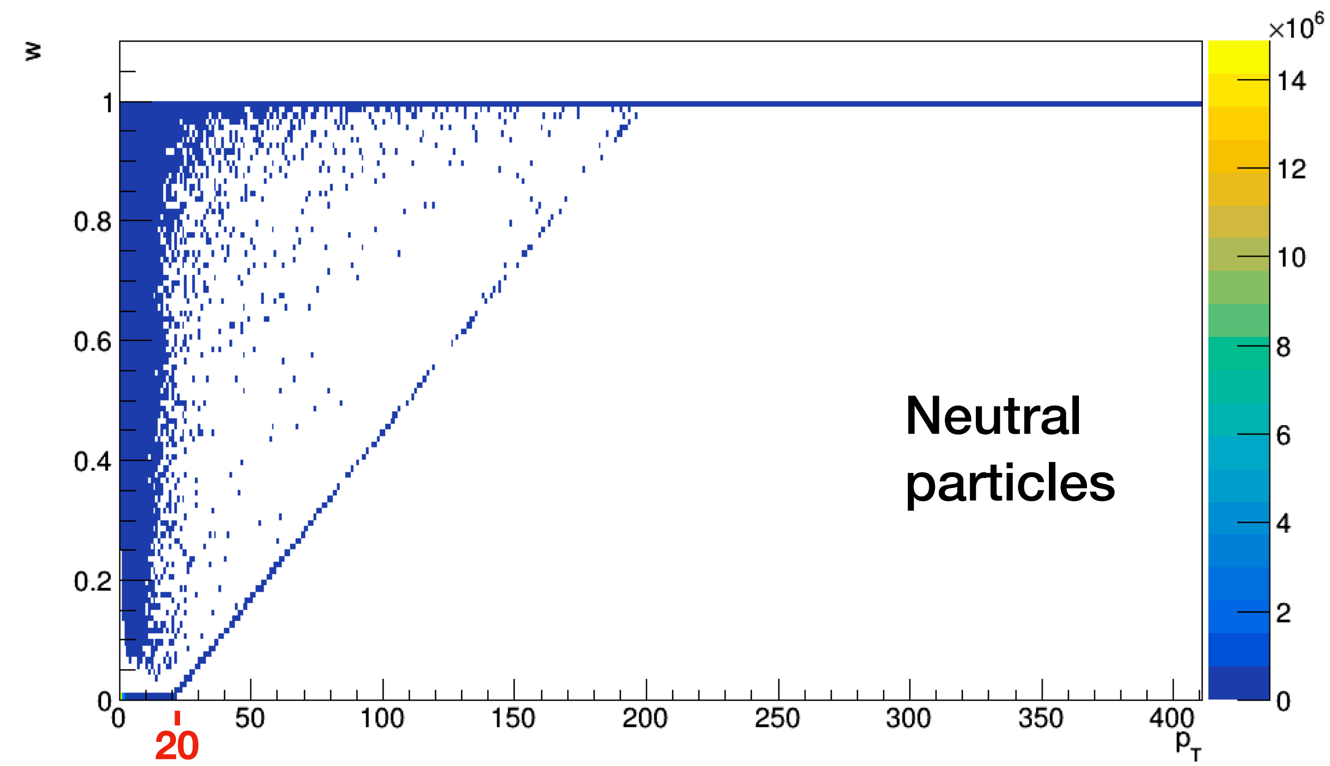
PUPPI tune v13 beagle

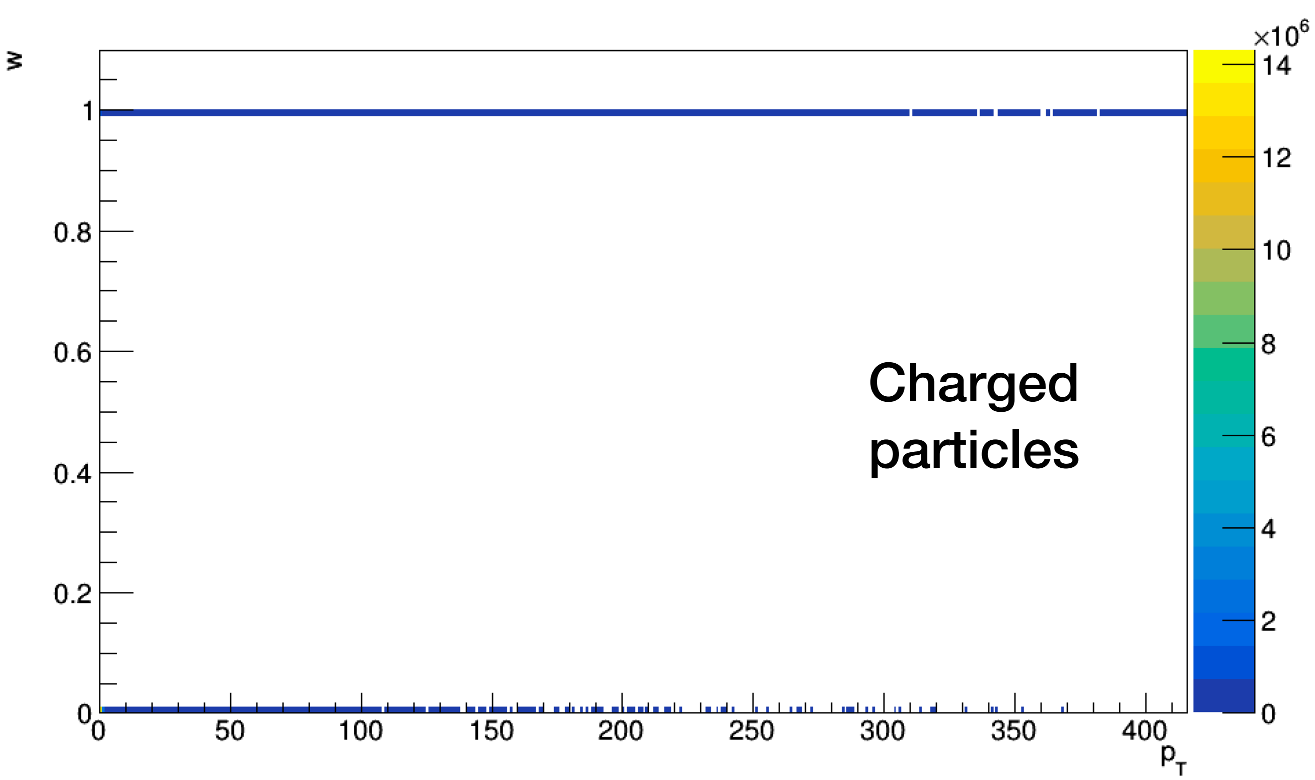
- **Wrong track-vertex association** at high p_T → remove d_Z cut and introduce p_T and η cuts
- Charged particles:
  - **p_T** protection at p_T = 20 GeV. **Below 20 GeV** particles are treated as neutrals ($0<w<1$), **above 20 GeV** all particles with no vertex assignment are kept
  - In 2017/2018 phase-1 pixel: more charged hadrons in **$|\eta|> 2.5$**, but **no track-vertex association** (fromPV=1 or 2)
  - Strategy: for $|\eta|> 2.5$ keep all the charged particles not associated to a PU vertex
- Neutral particles:
  - Linear **p_T** protection between 20 and 200 GeV

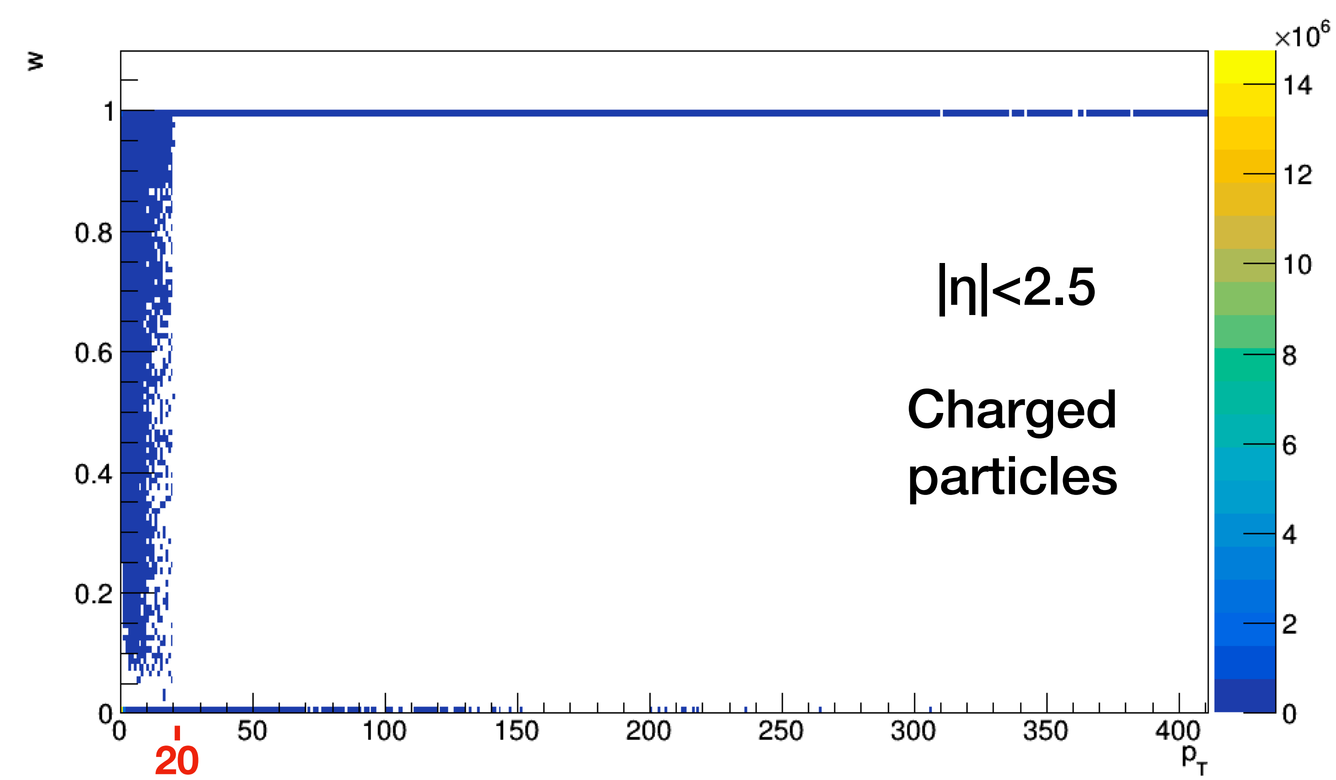
weight vs p_T

PUPPI default | **PUPPI v13 beagle**





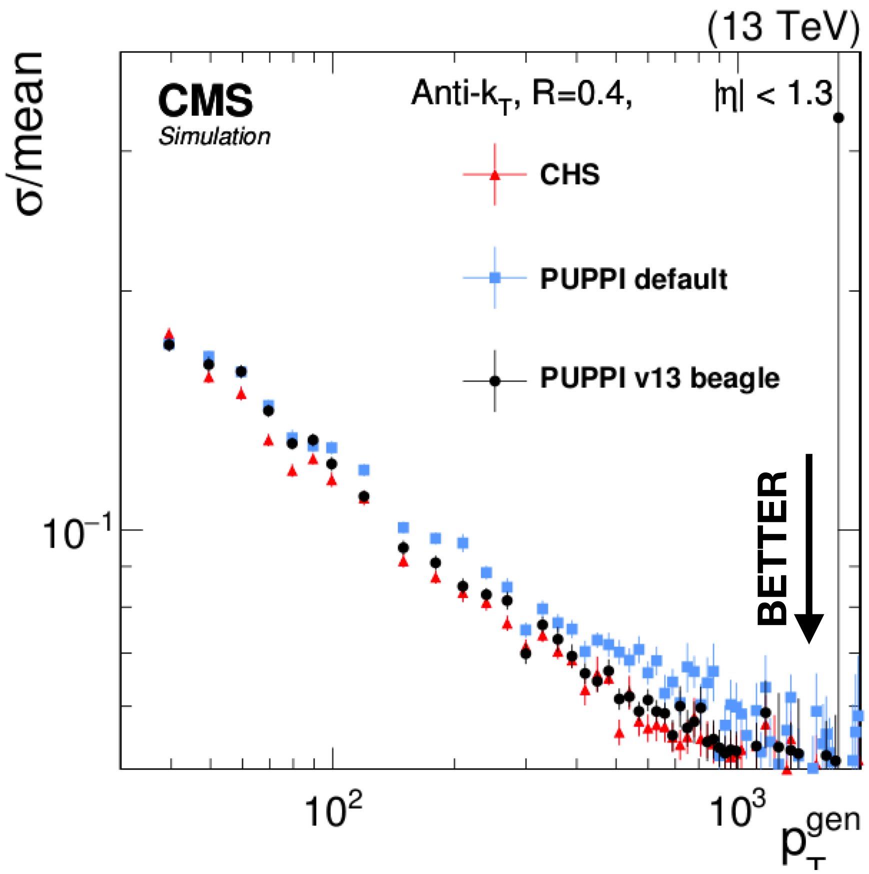


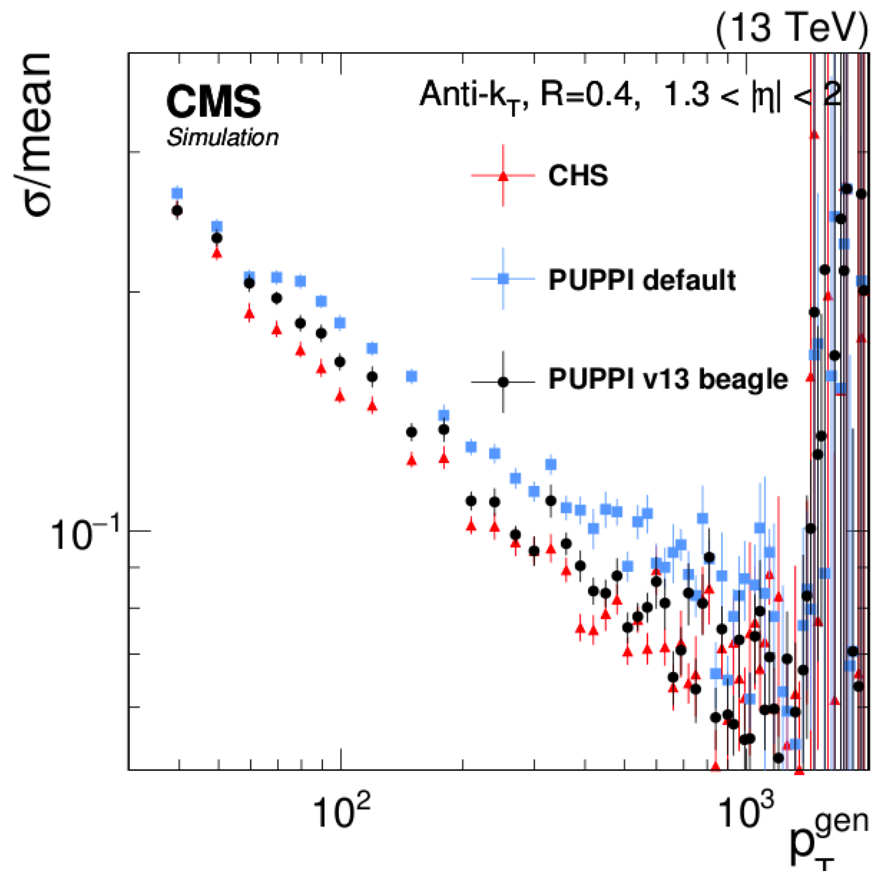


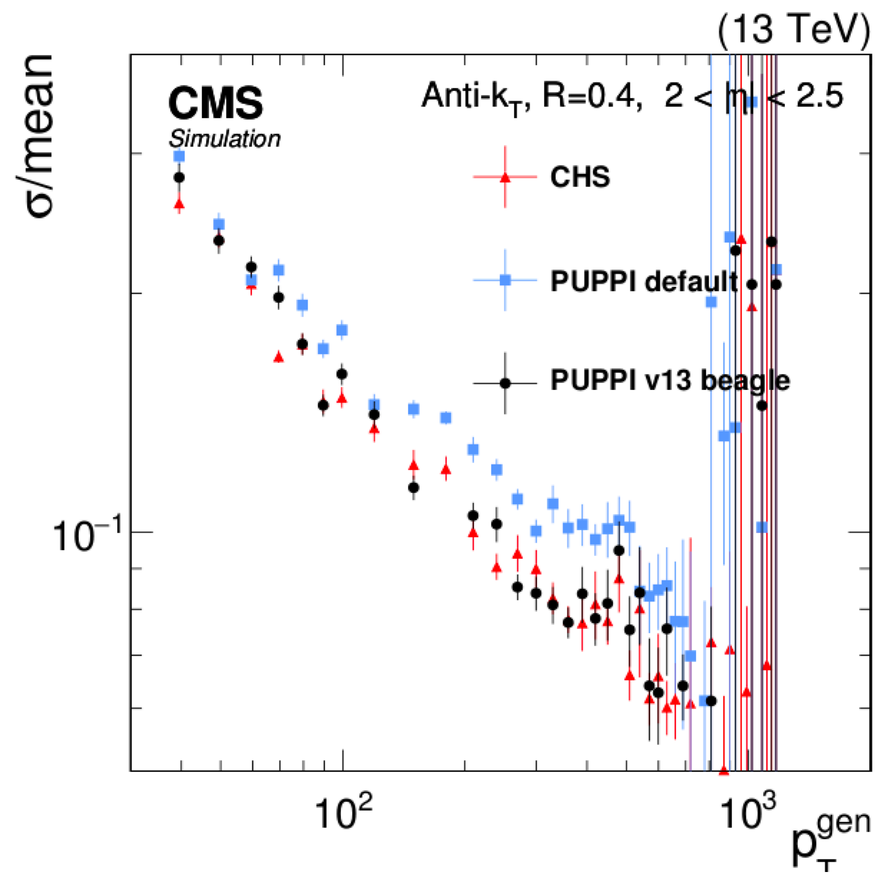
sample: QCD 2018

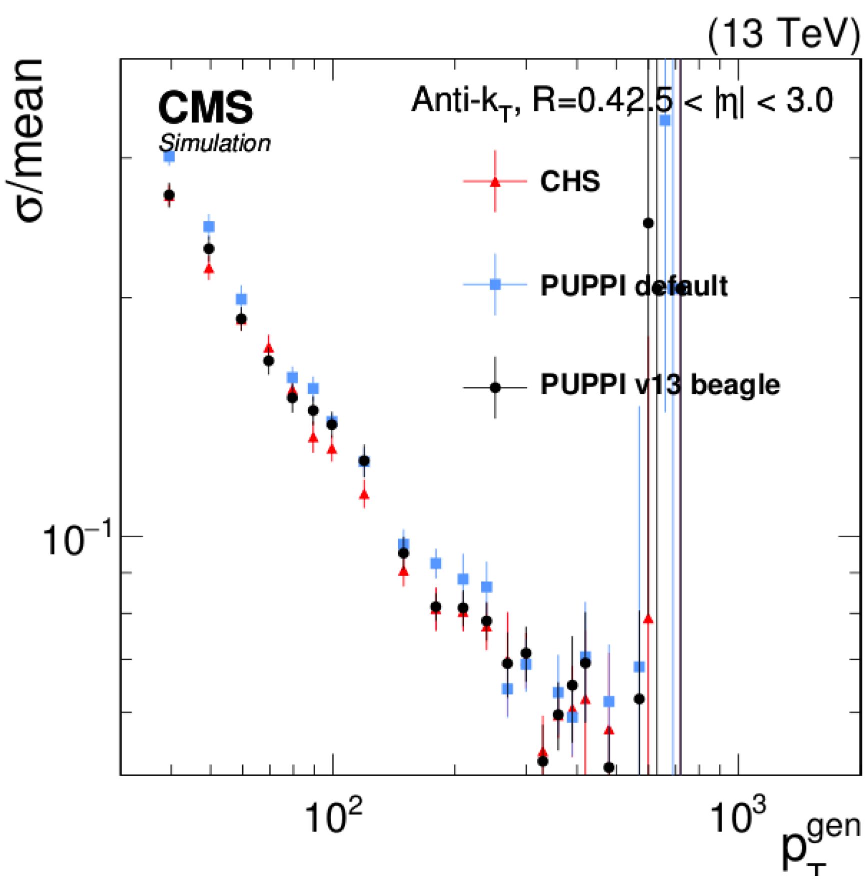
JER 2017 - Ak4 PU>50

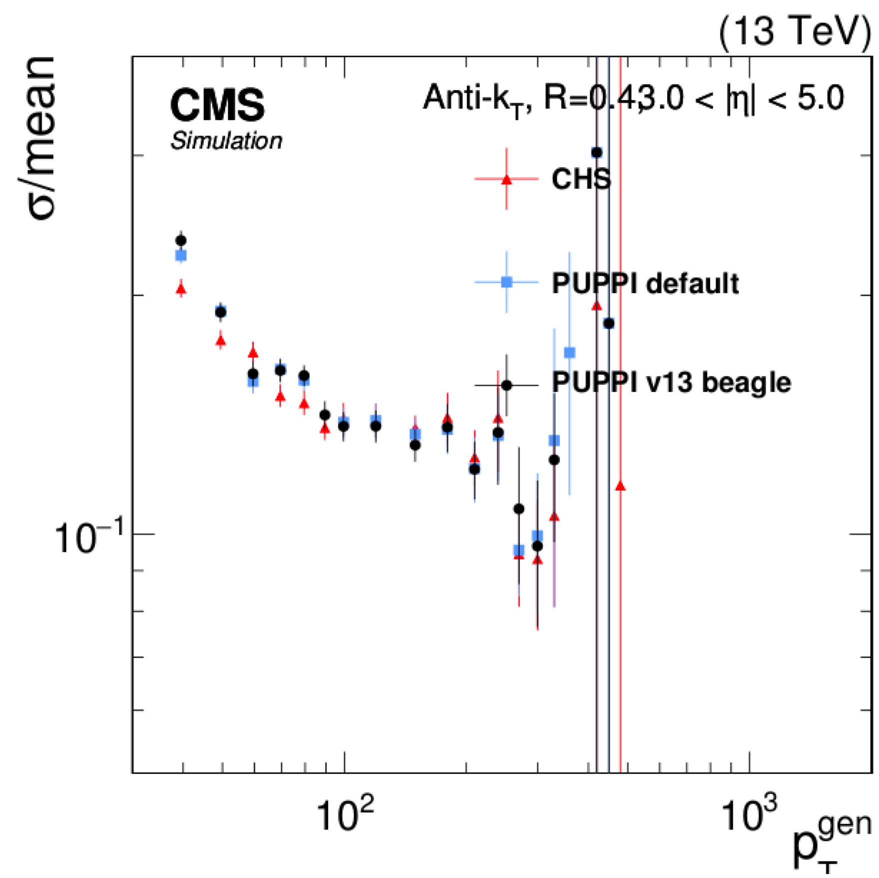
QCD sample







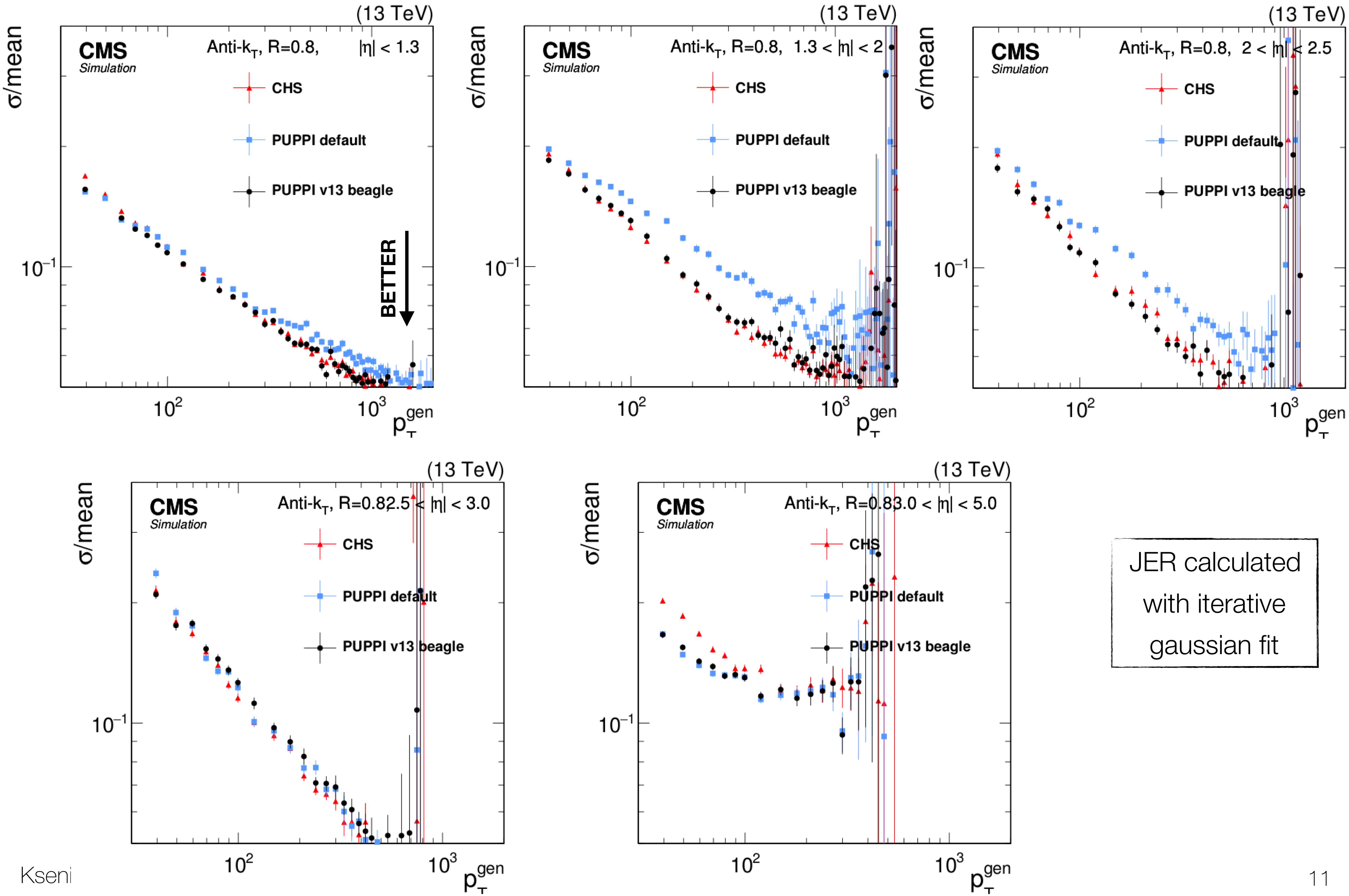




JER calculated with iterative gaussian fit

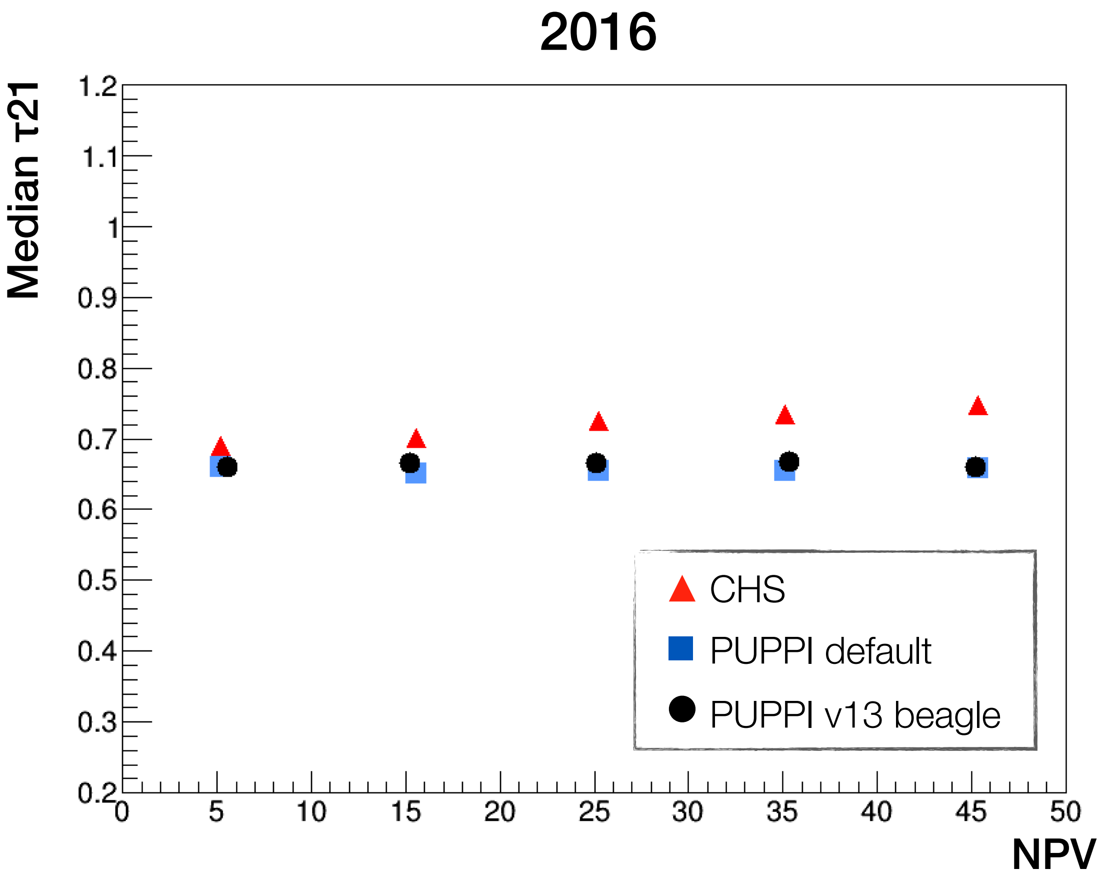
JER 2017 - Ak8

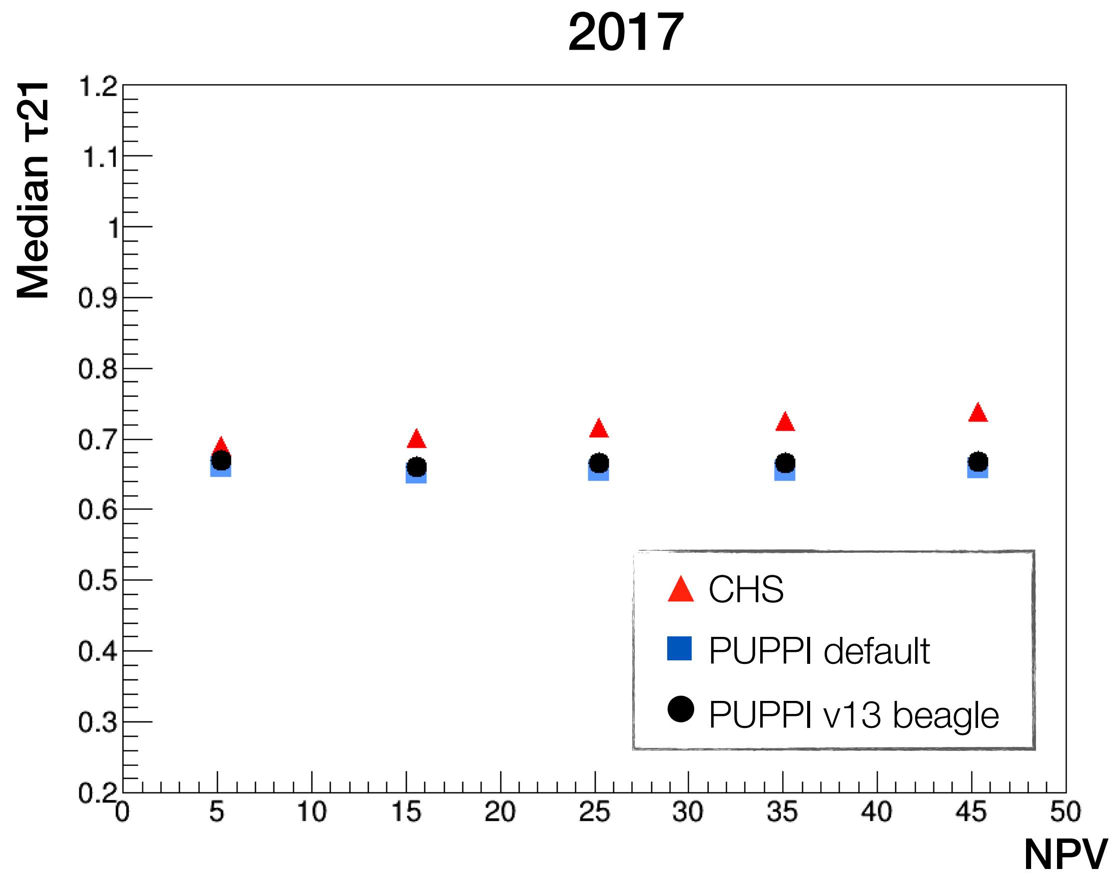
QCD sample

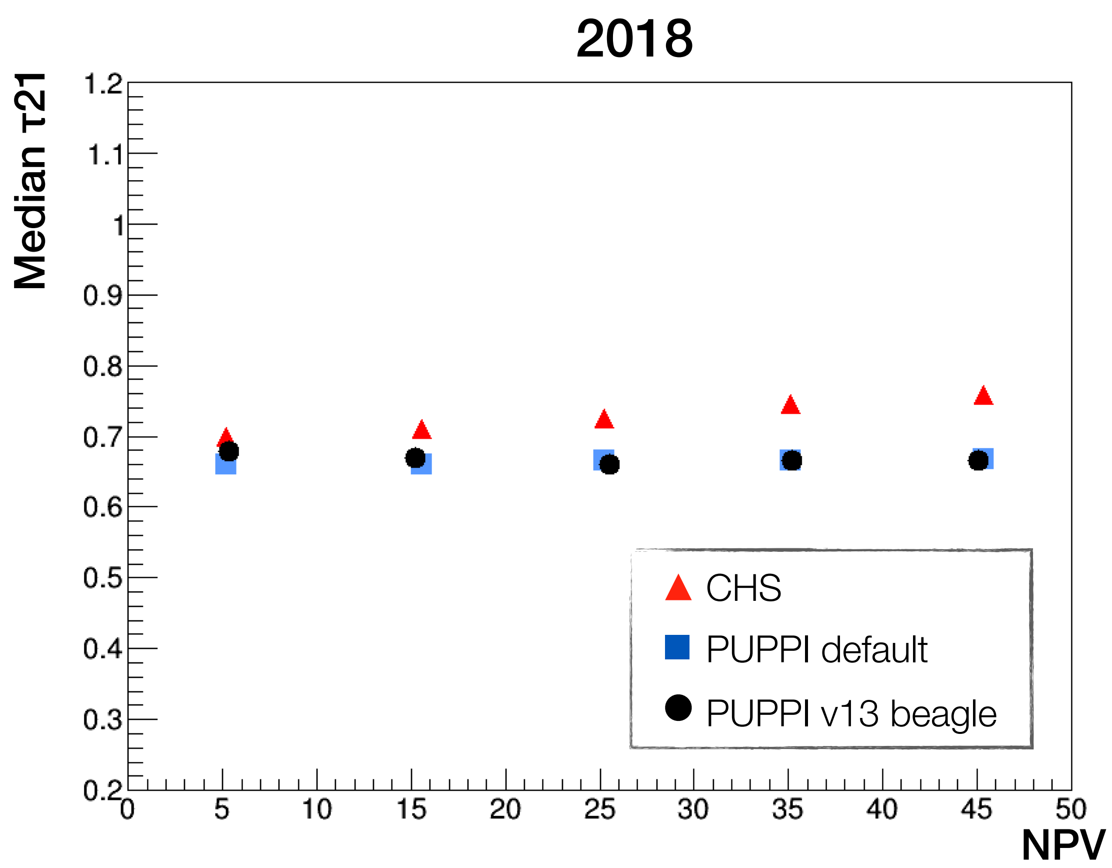


JER calculated with iterative gaussian fit

Median τ21







QCD sample

Median τ21 leading Ak8 jet

400 < p_T < 600 GeV

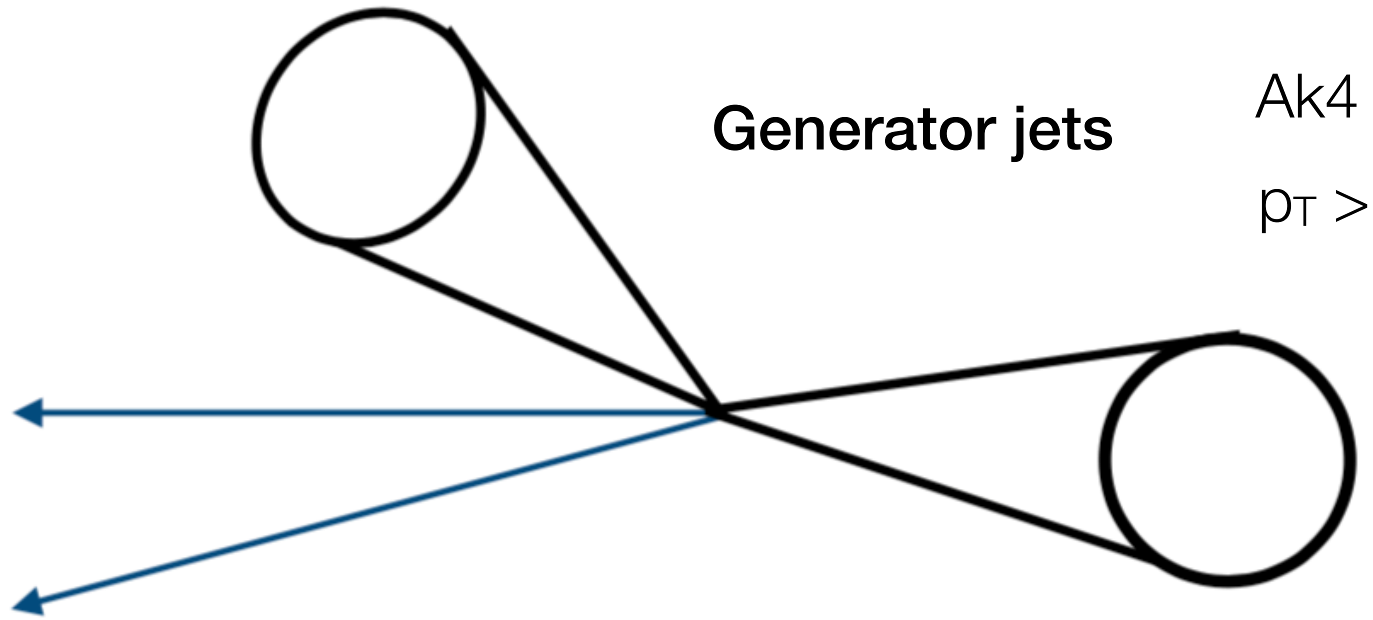
no SoftDrop Mass cut

Efficiency & Purity

Efficiency and Purity

Z + Jets

MC studies



Generator jets

Ak4 jets with

$p_T > 20$ GeV

muons

2 muons with $p_T > 20$ GeV,

70 GeV < $M_{\mu\mu}$ < 110 GeV

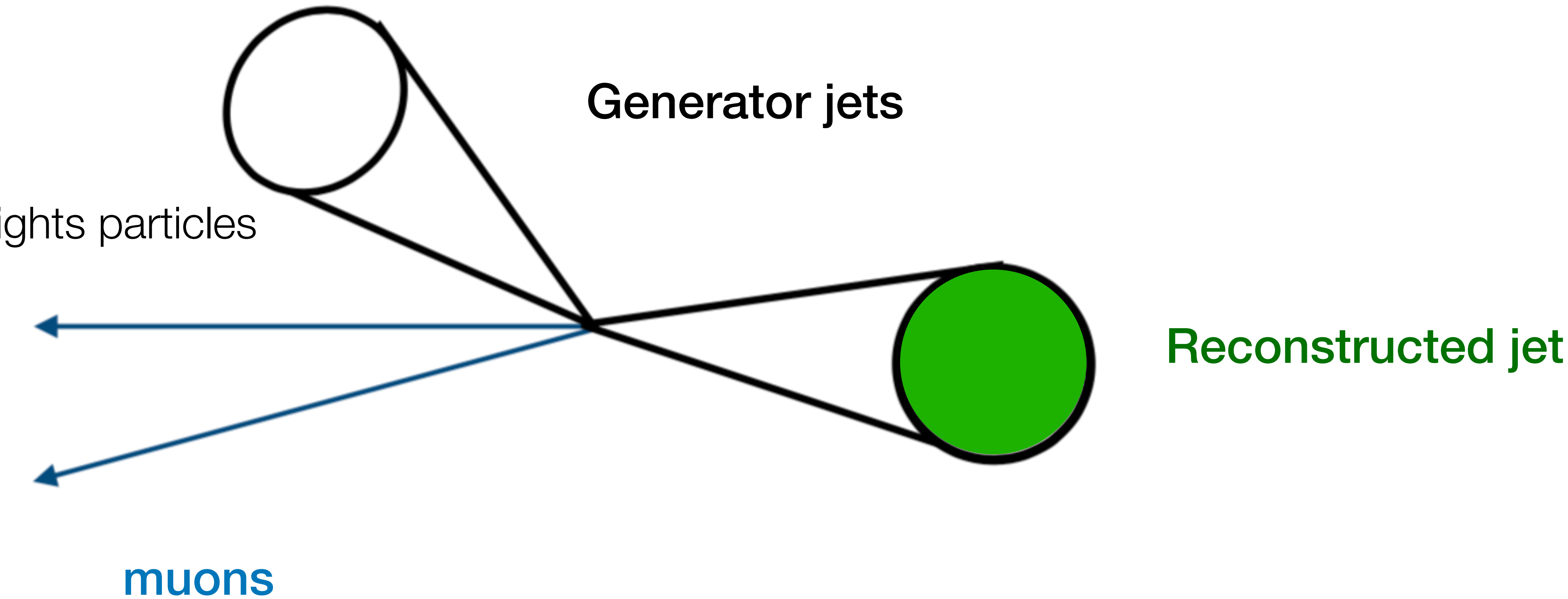
Jets overlapping with muons removed

Efficiency and Purity

Efficiency

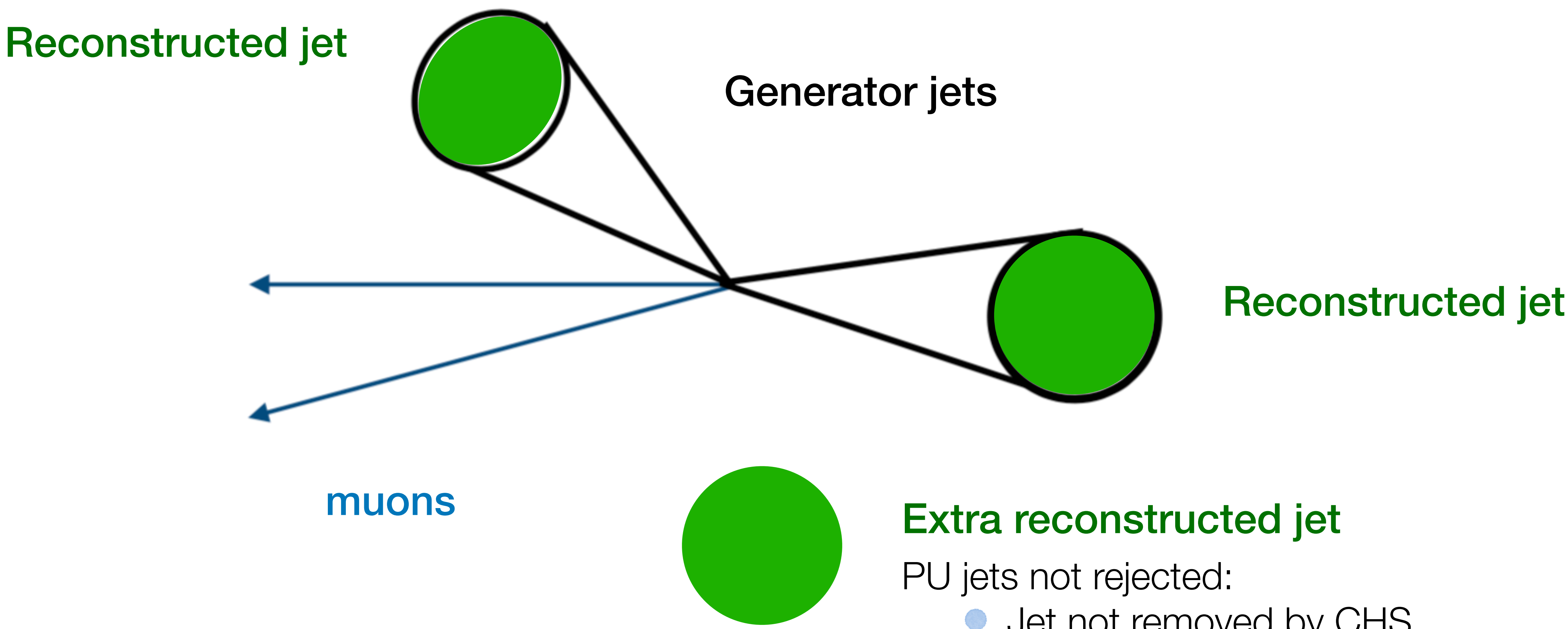
Not reconstructed or rejected jet

- Pileup Jet ID
- PUPPI down-weights particles



Efficiency and Purity

Purity

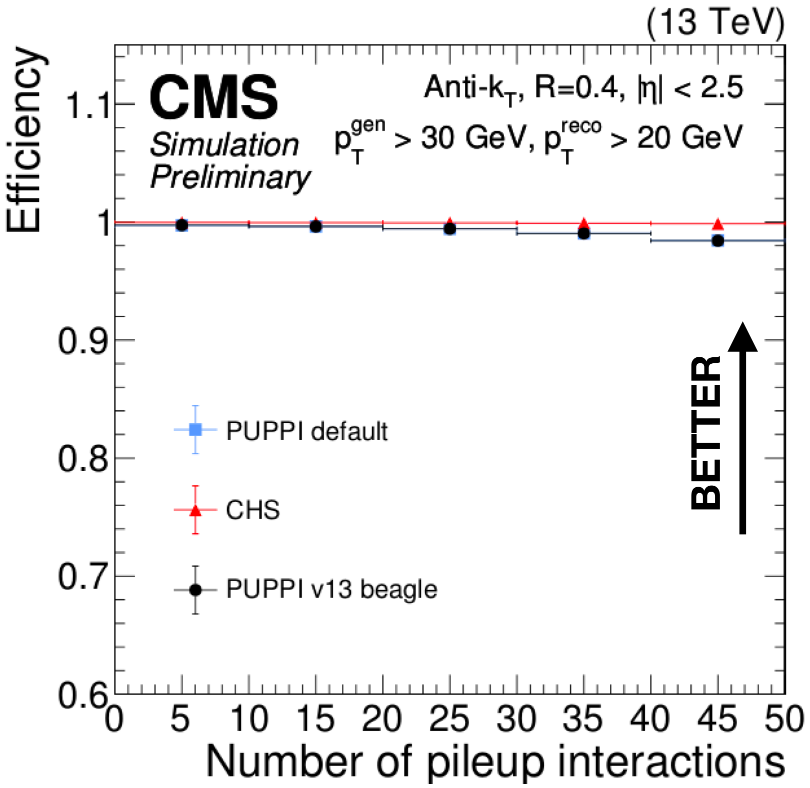


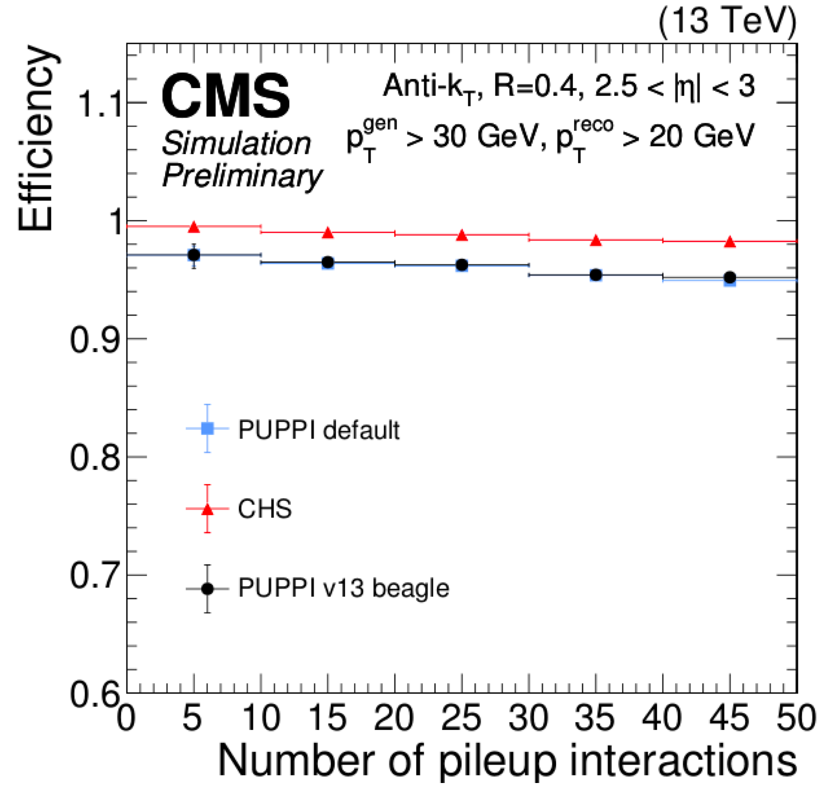
PU jets not rejected:

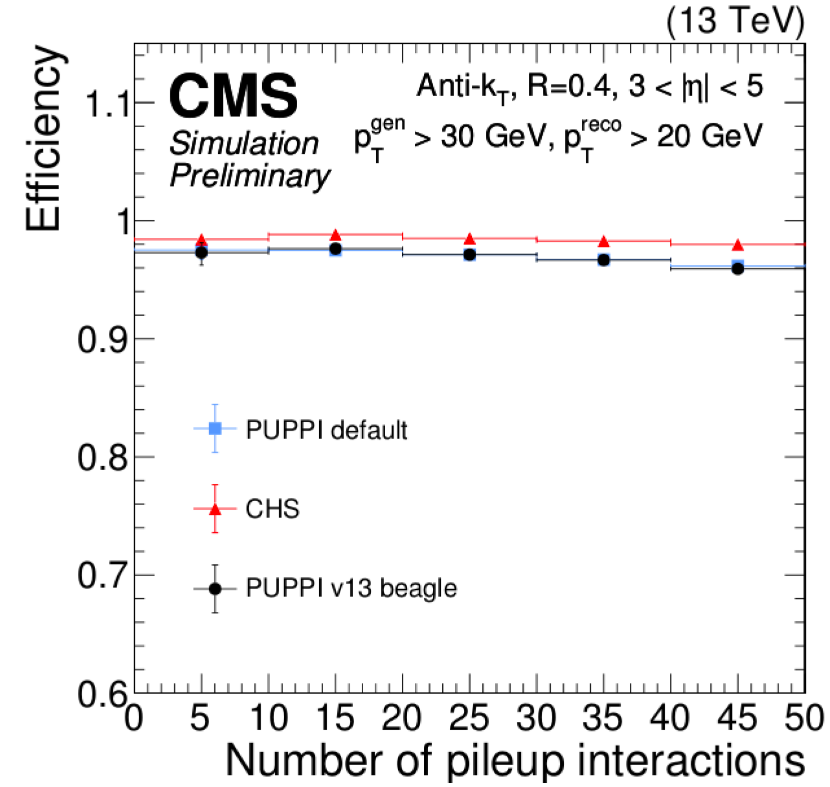
- Jet not removed by CHS
- Particles not removed by PUPPI

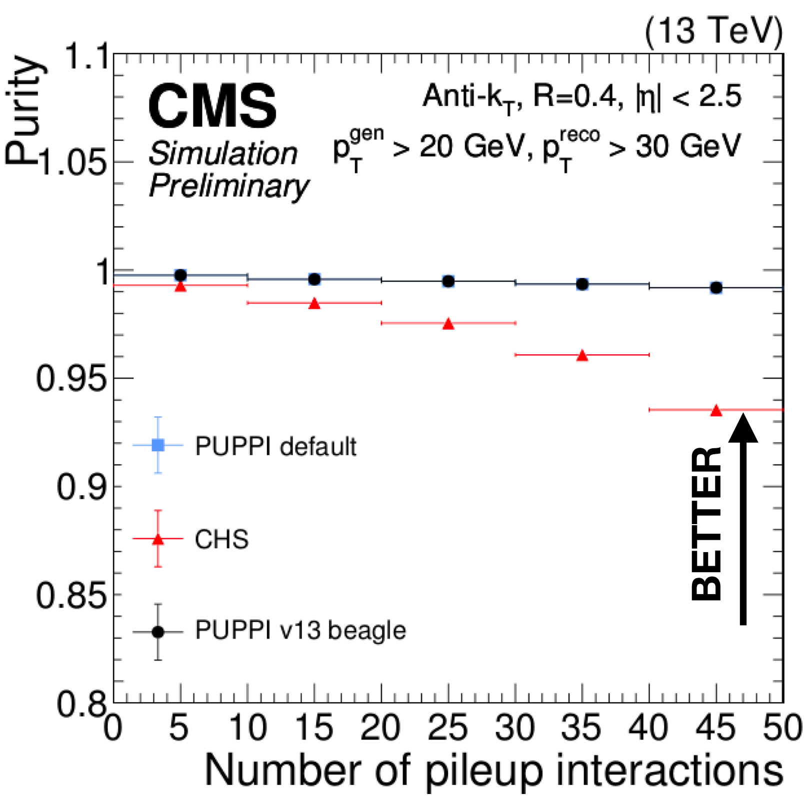
Efficiency & Purity - 2016

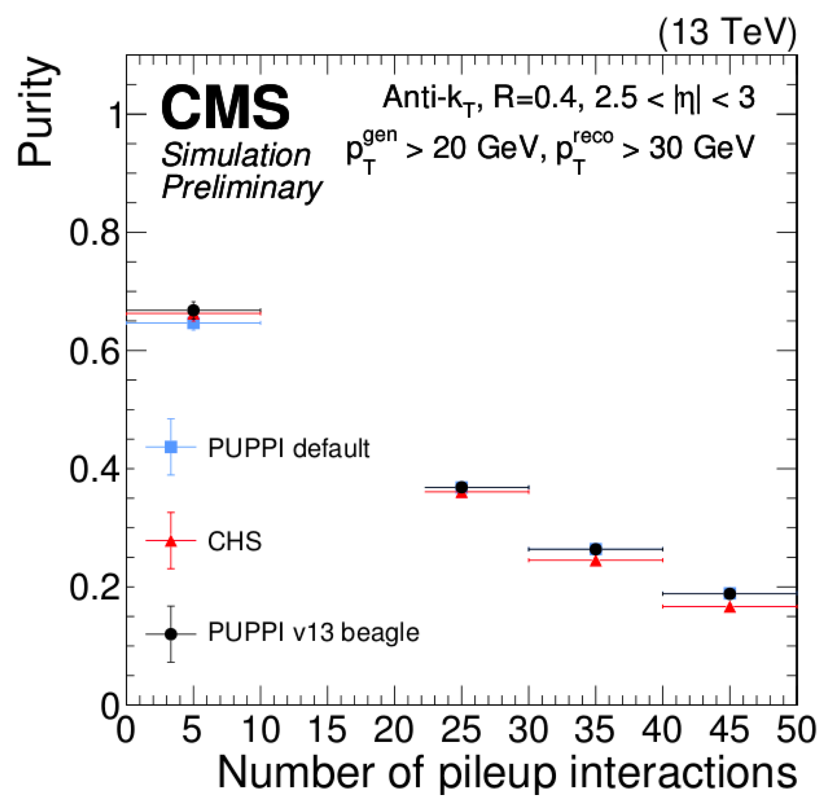
DY sample

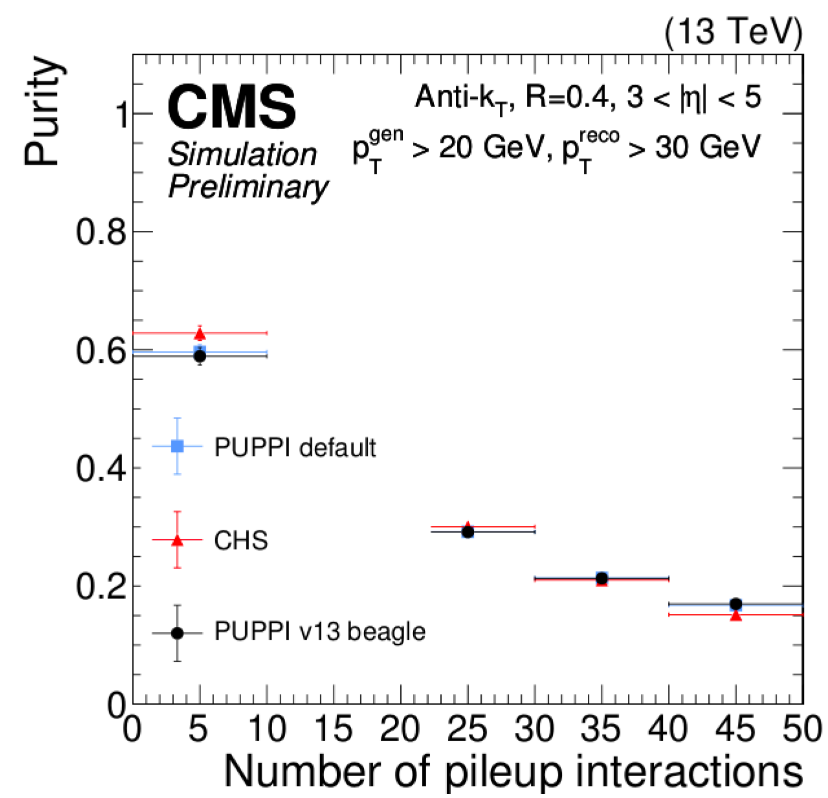






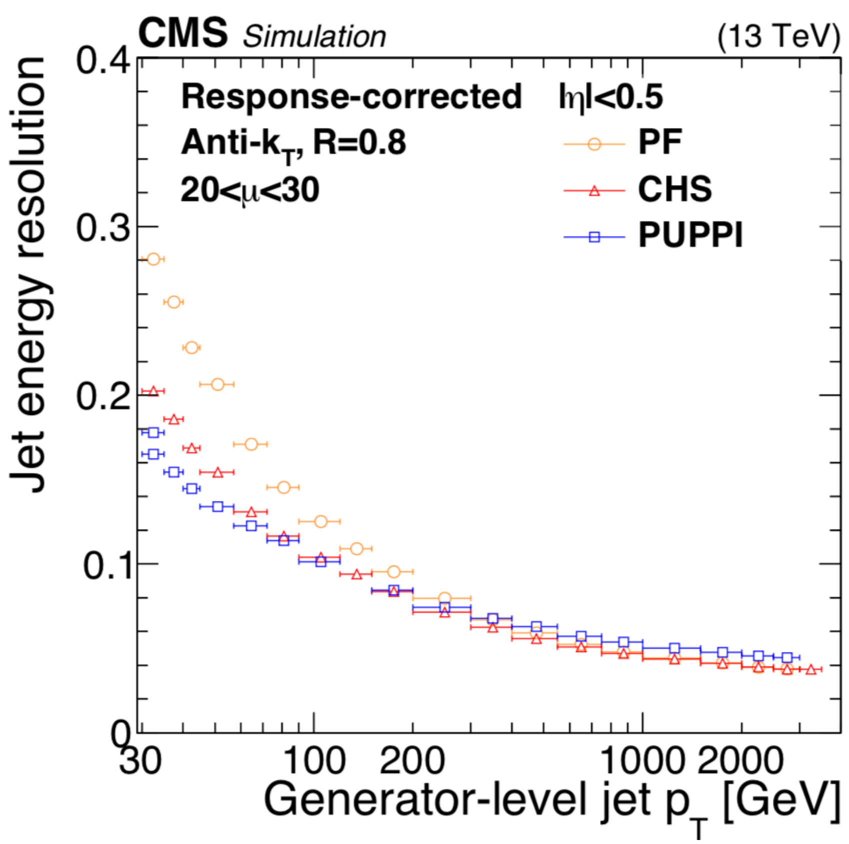


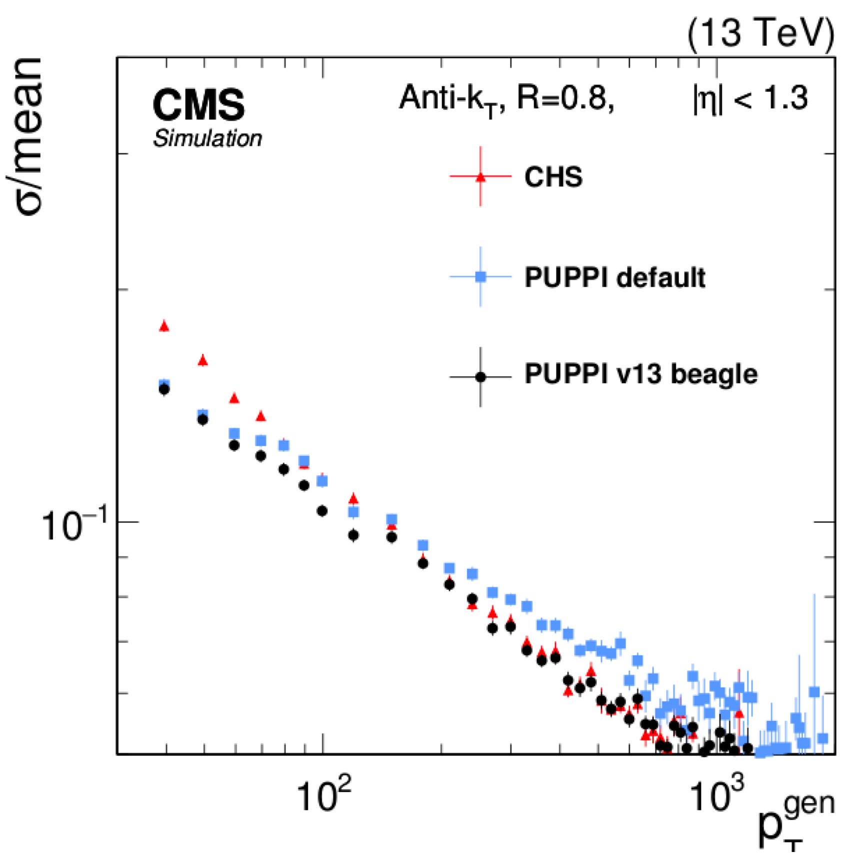




Summary

- Pileup mitigation in jets
- PUPPI shows very good performance → standard algorithm in Run3
- Improvements: **JER** at high p_T
- New tune **PUPPI v13 beagle** and preliminary validation





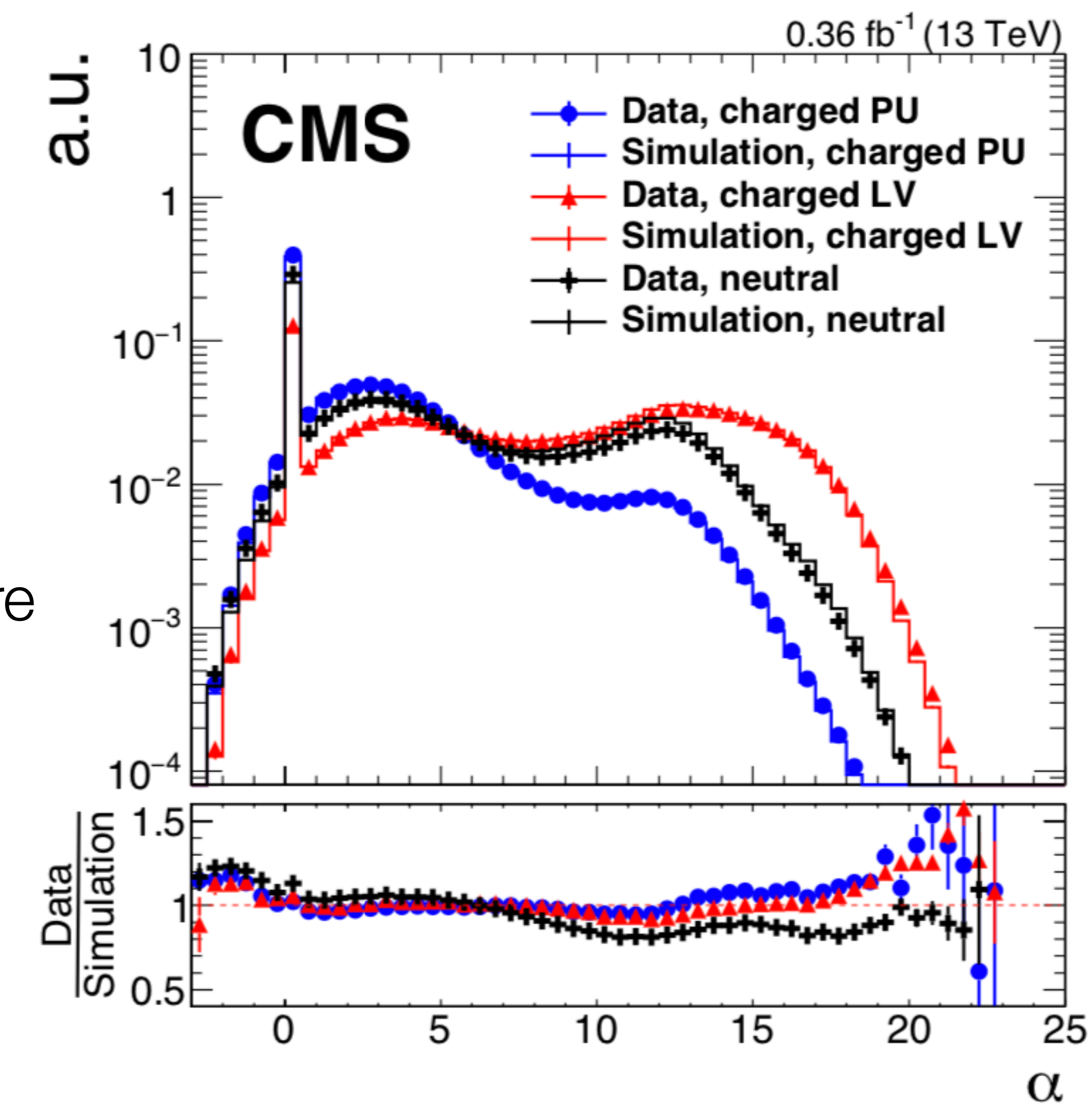
BACKUP

PUPPI details - 1

- Define α variable for neutral particle *i*:

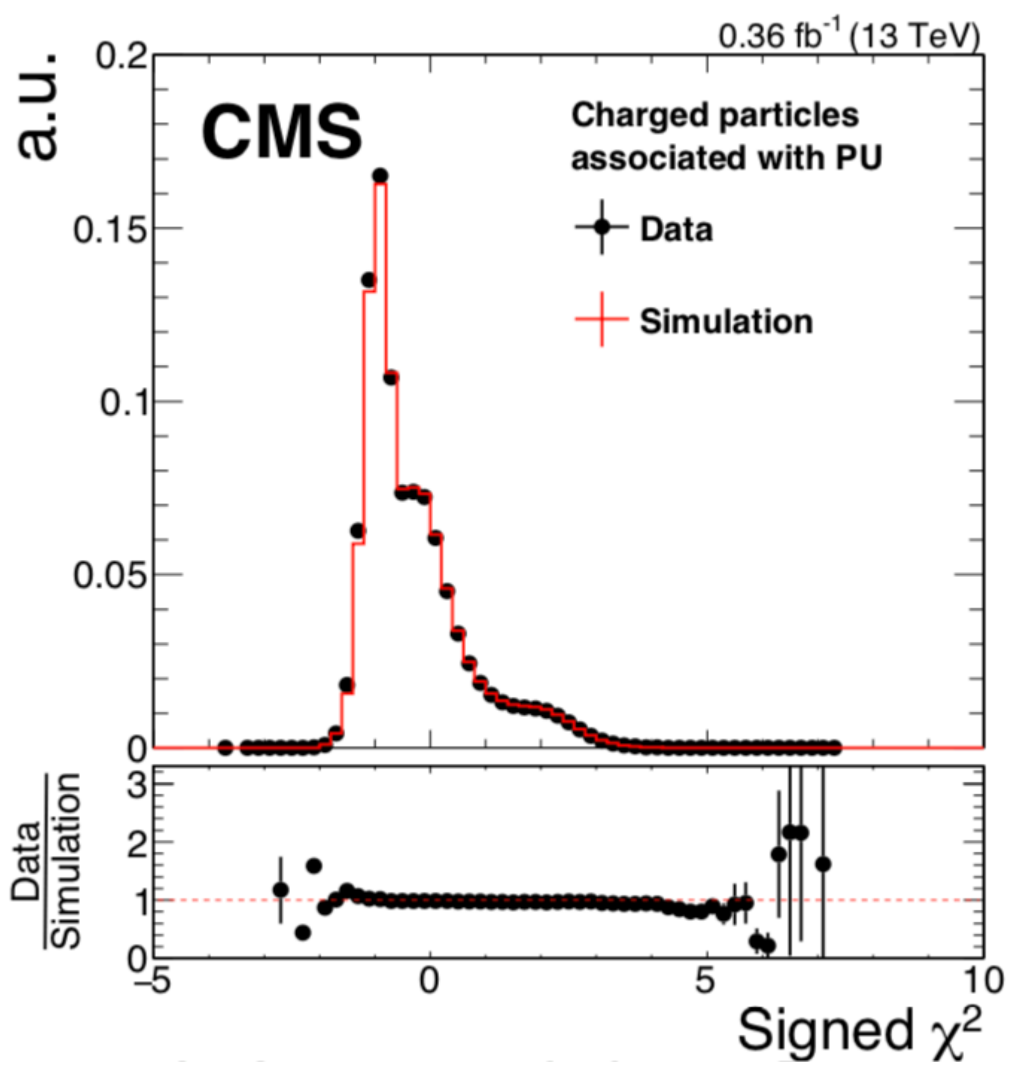
$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{Tj}}{\Delta R_{ij}} \right)^2 \begin{cases} \text{for } |\eta_i| < 2.5, & j \text{ are charged particles from LV,} \\ \text{for } |\eta_i| > 2.5, & j \text{ are all kinds of reconstructed particles} \end{cases}$$

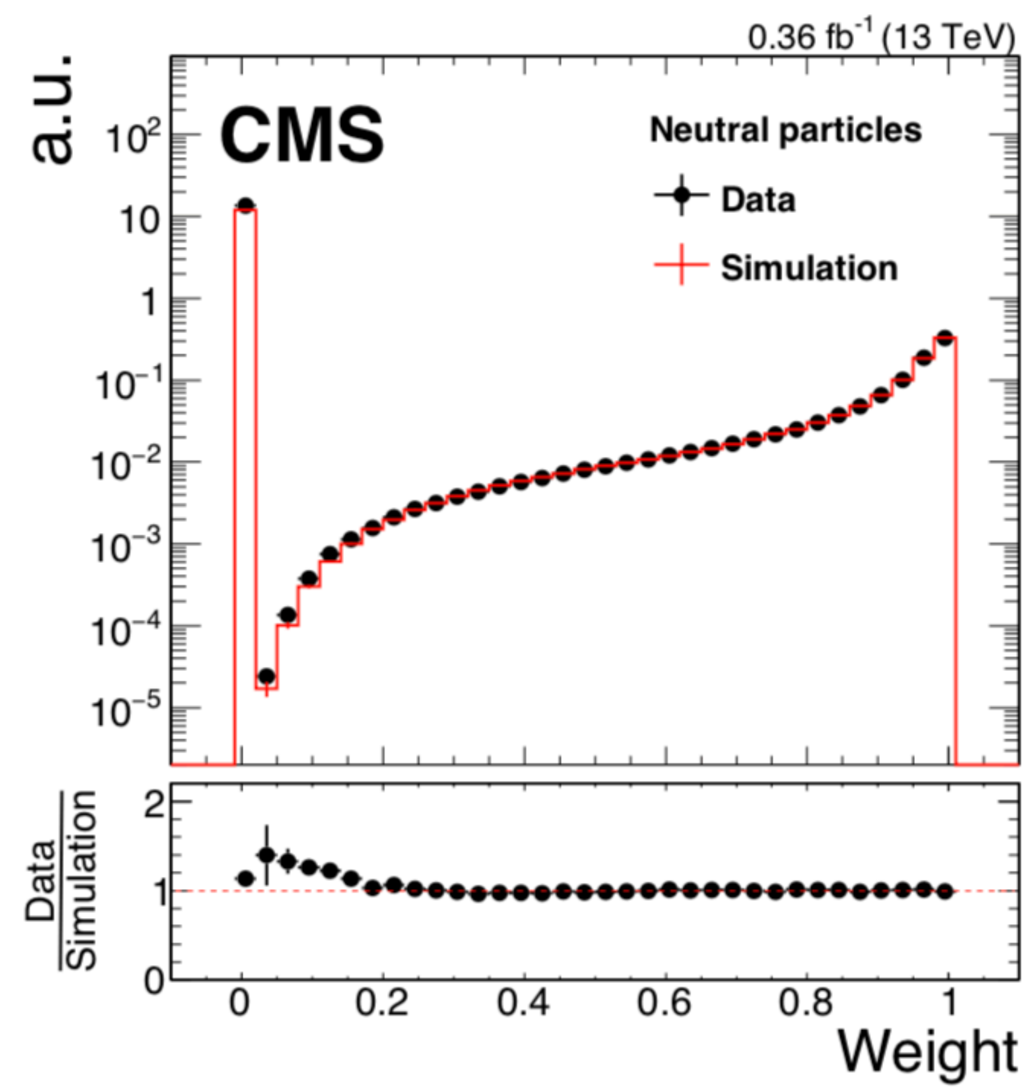
- Suppose neutral PU same shape as charged PU
- Calculate α distribution of charged PU, extract median $\overline{\alpha}_{PU}$ and root-mean-square RMS_{PU} of the distribution



PUPPI details - 2

- For each particle *i* calculate: $$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})|\alpha_i - \bar{\alpha}_{PU}|}{RMS_{PU}^2}$$
- Assign a weight w_i: $$w_i = F_{\chi^2, NDF=1}(\chi_i^2)$$





PUPPI details - 3

- Remove particles with $w_i < 0.01$
- Keep only neutral particles that fulfil:

$$w_i \times p_{T,i} > (A + B \times NPV) \text{ GeV}$$

- A, B: tunable parameters
- NPV: number of primary vertices

How PUPPI can treat Charged Particles

Properties	from LV	from PU	α_{PU} calculation	α_i calculation	weight	note
1	✓			✓	1	Charged LV
2		✓	✓		0	Charged PU
0				✓	0 to 1	Neutrals

Tune v13

- In 2017 and 2018 the **v13 beagle** tune shows a **worse** performance in the region **|η|> 2.5**
- More charged hadrons in |η|> 2.5, but **no track-vertex association**: the particles get a value of fromPV=1 or 2
- Strategy: in |η|> 2.5 keep all the charged particles with fromPV=1,2,3

pT cut	**pT < 20 GeV**		**pT > 20 GeV**	
eta cut	**|η| < 2.5**	**|η| > 2.5**	**|η| < 2.5**	**|η| > 2.5**
v13 beagle	**N**	**N**	**LV**	**LV**
v13 new PR 2017	**N**	**LV**	**LV**	**LV**

A charged particle not associated to any vertex can be treated as:
1 Charged LV (LV) 2 Charged PU (PU) 0 Neutral (N)

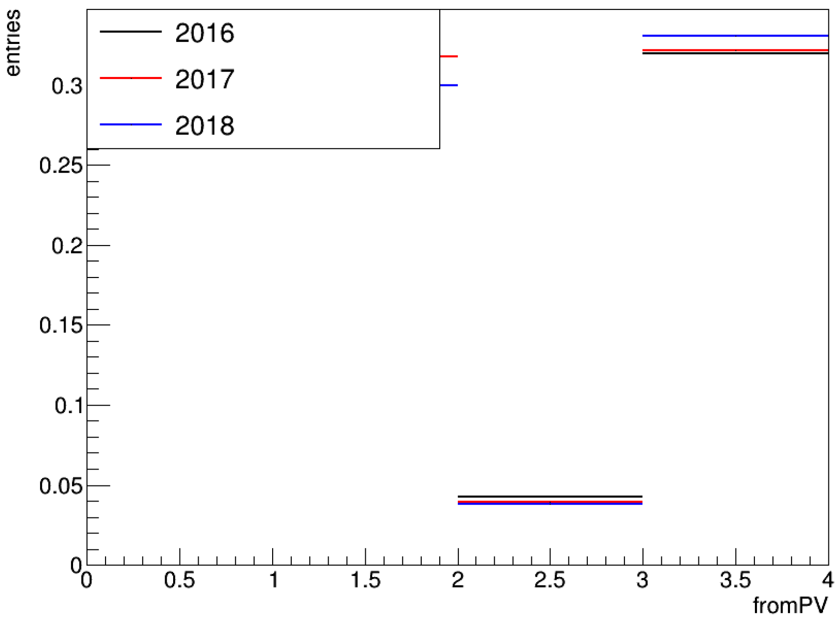
QCD Dataset

Year	DAS	CMSSW release	Notes
2016v2	/QCD_Pt-15to7000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIISummer16MiniAODv2-PUFlat0to70_magnetOn_80X_mcRun2_asymptotic_2016_TrancheIV_v4-v1/MINIAODSIM	production: 80X reconstruction: 80X	
2017v2	/QCD_Pt-15to7000_TuneCUETP8M1_Flat_13TeV_pythia8/RunIIFall17MiniAODv2-FlatPU0to70_12Apr2018_94X_mc2017_realistic_v14-v1/MINIAODSIM	production: 92X reconstruction: 94X	• More particles at low p_T • Phase-1 pixel: more charged particles in $\|\eta\| > 2.5$
2018	/QCD_Pt-15to7000_TuneCP5_Flat_13TeV_pythia8/RunIIAutumn18MiniAOD-FlatPU0to70RAW_102X_upgrade2018_realistic_v15_ext2-v1/MINIAODSIM	production: 102X reconstruction: 102X	• More particles at low p_T • Phase-1 pixel: more charged particles in $\|\eta\| > 2.5$
UL2017	/QCD_Pt-15to7000_TuneCP5_Flat2018_13TeV_pythia8/RunIISummer19UL17MiniAOD-FlatPU0to70_106X_mc2017_realistic_v6-v3/MINIAODSIM	production: 106X reconstruction:106X	Coming soon
UL2018	/QCD_Pt-15to7000_TuneCP5_Flat2018_13TeV_pythia8/RunIISummer19UL18MiniAOD-pilot_106X_upgrade2018_realistic_v11_L1v1-v3/MINIAODSIM	production: 106X reconstruction:106X	Coming soon

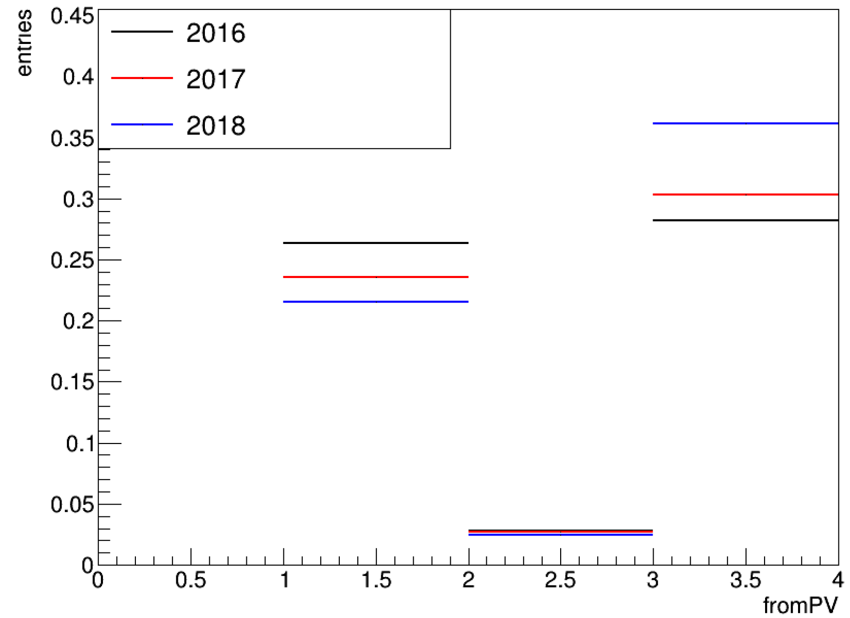
PR tune v13 beagle

fromPV - p_T < 20 GeV

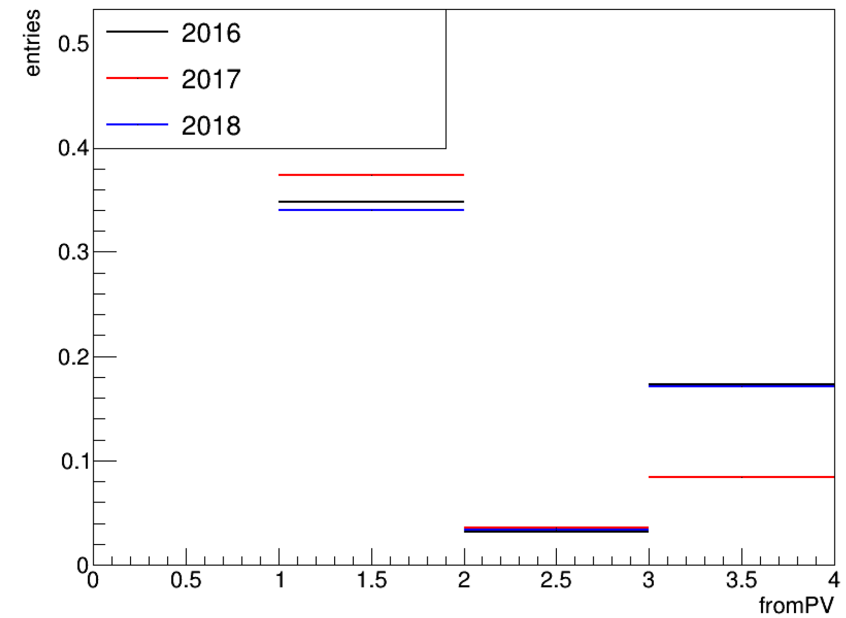
0 < |η| < 1.3



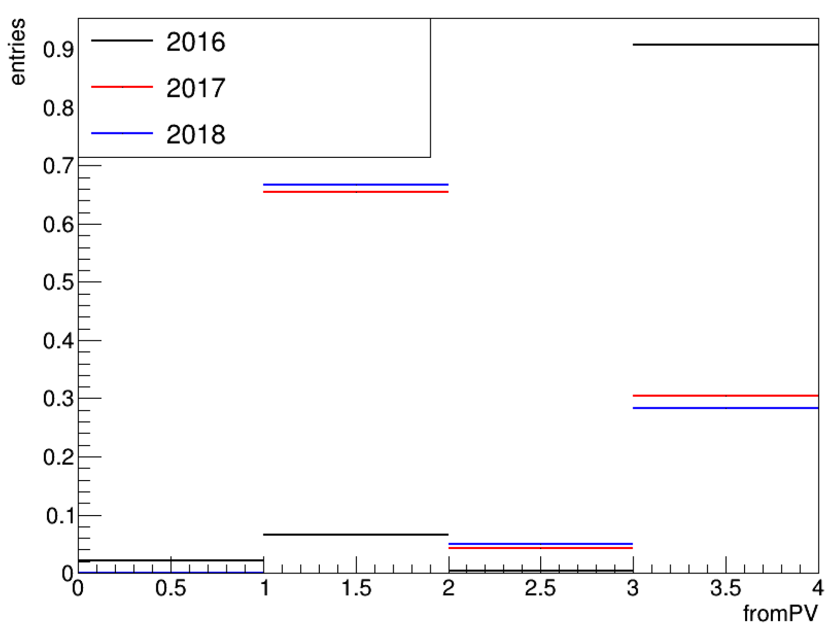
1.3 < |η| < 2



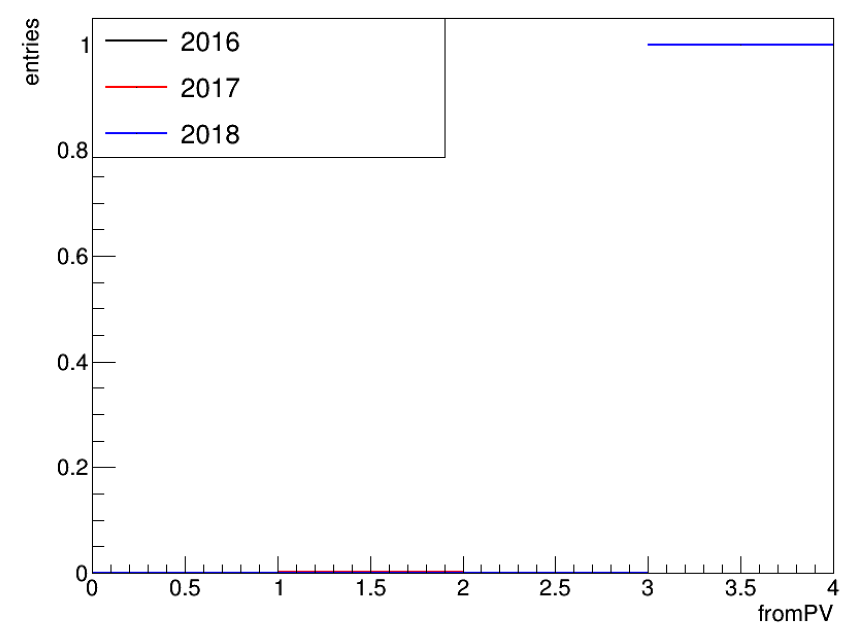
2 < |η| < 2.5



2.5 < |η| < 3

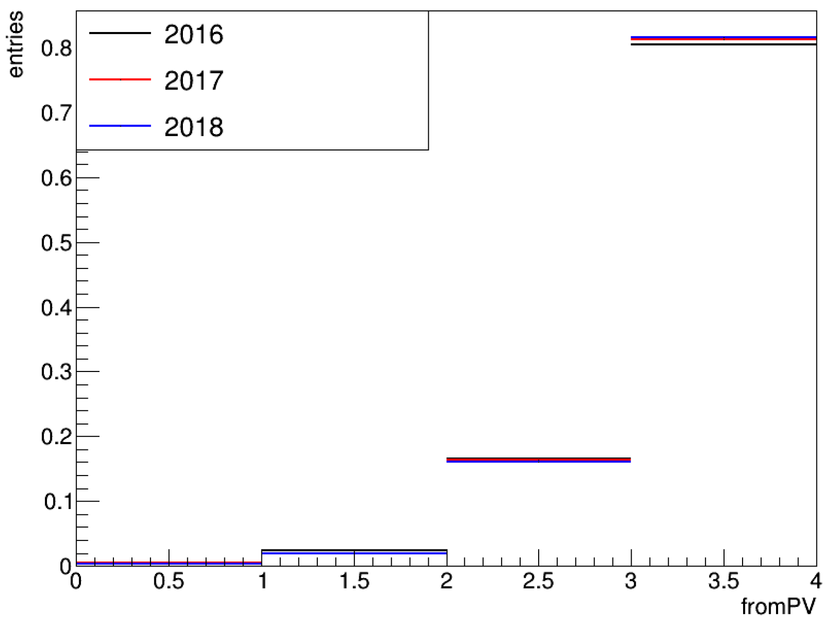


3 < |η| < 10

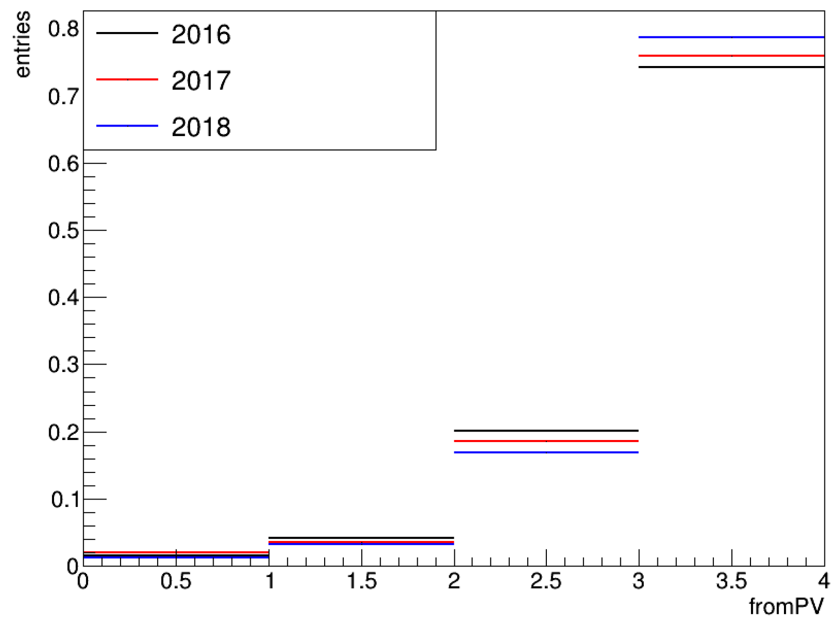


fromPV - p_T > 20 GeV

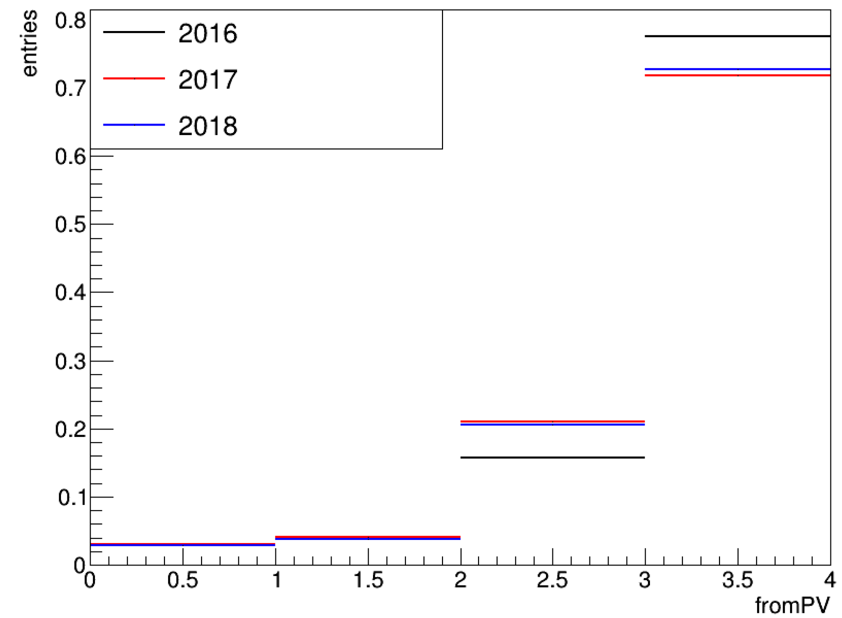
0 < |η| < 1.3



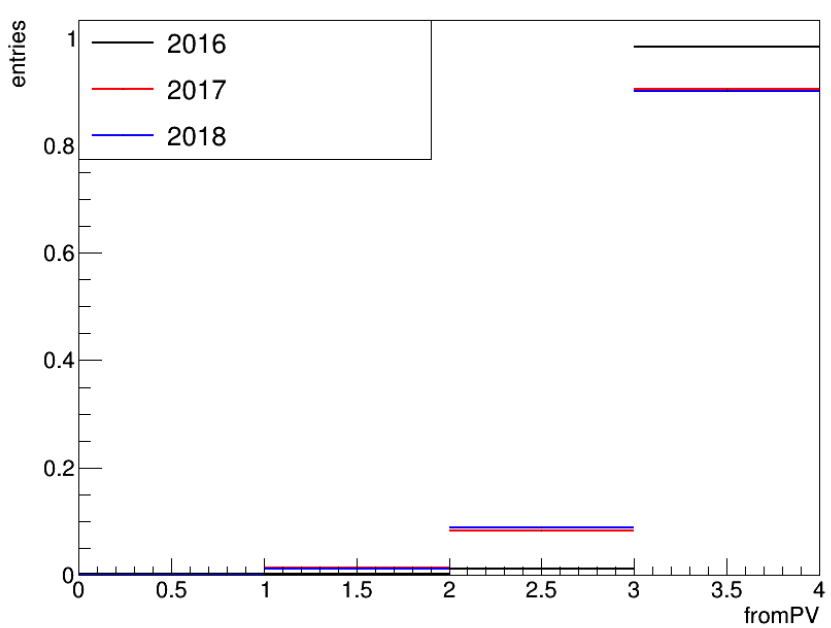
1.3 < |η| < 2



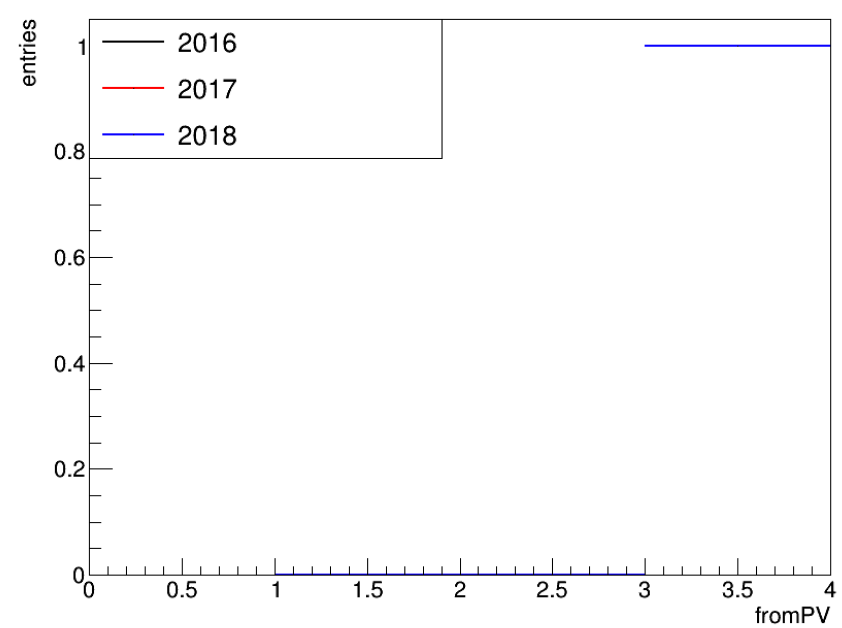
2 < |η| < 2.5



2.5 < |η| < 3

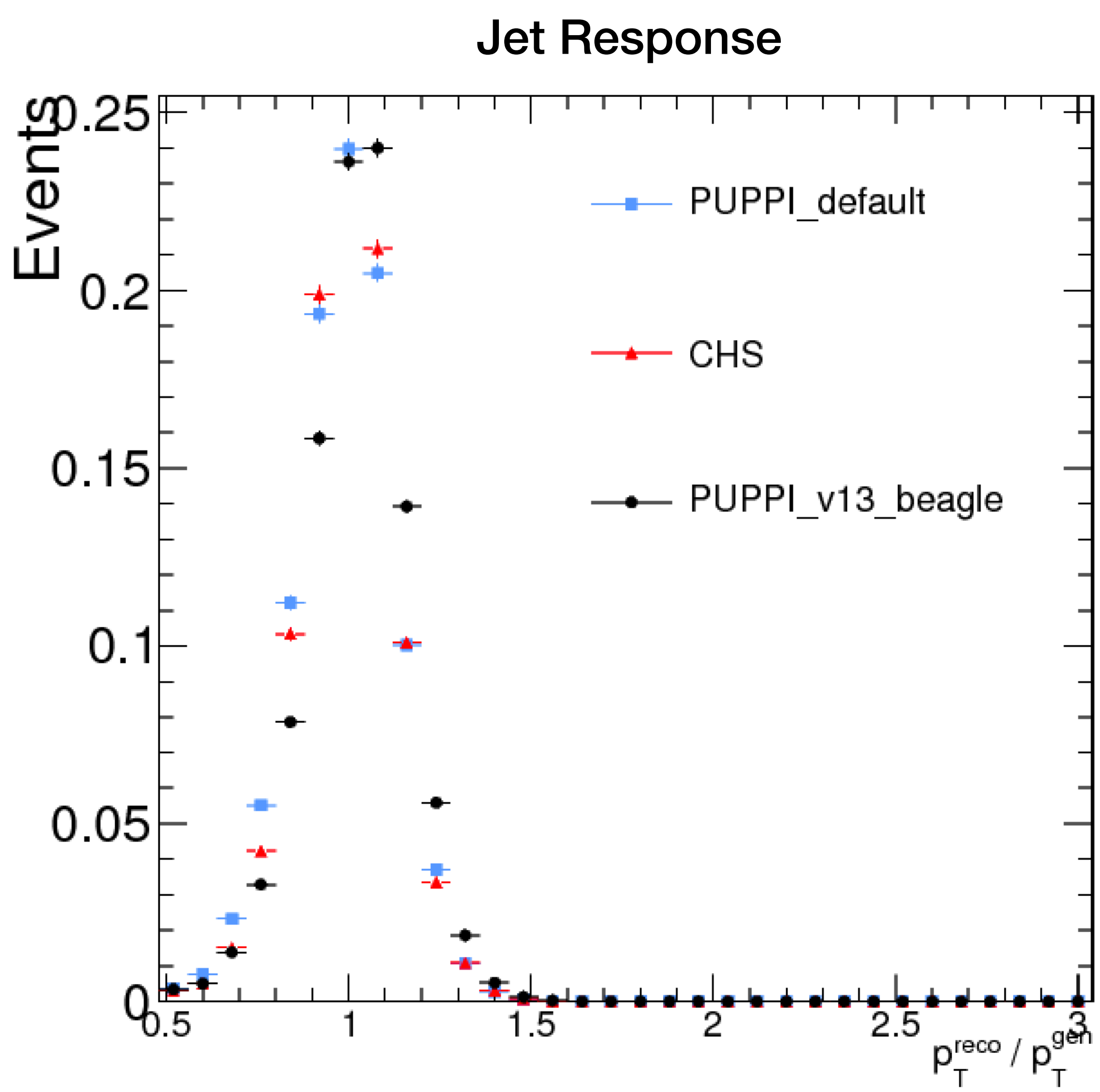


3 < |η| < 10



JER calculation

- Response distribution: p_T $^{(recojets)}$/p_T $^{(genjets)}$
- Iterative Gaussian fit in the range [*mean* - 1.5σ, *mean* + 1.5σ]
- JER = σ/*mean*

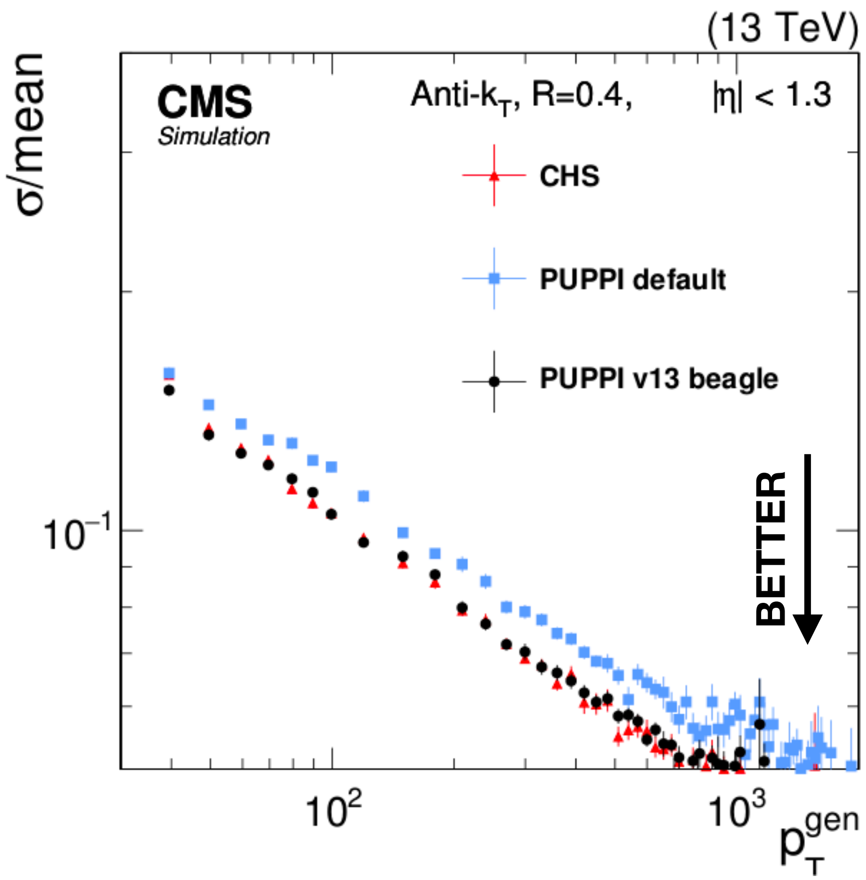


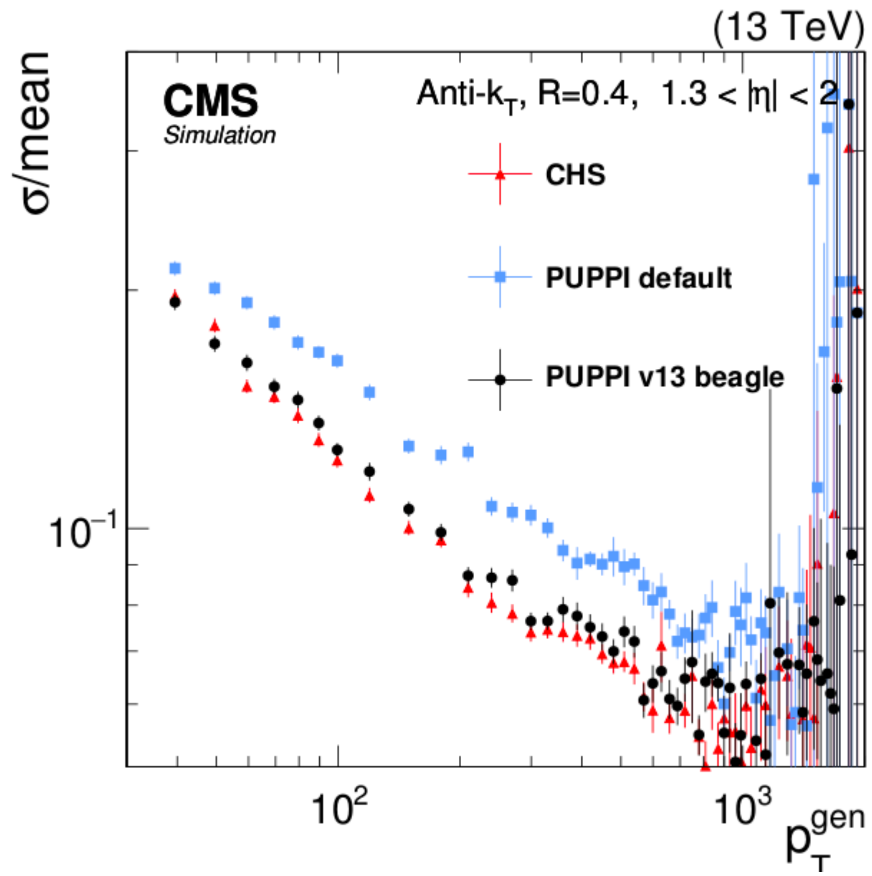
0 < |η| < 1.3
80 < jet p_T < 90 GeV

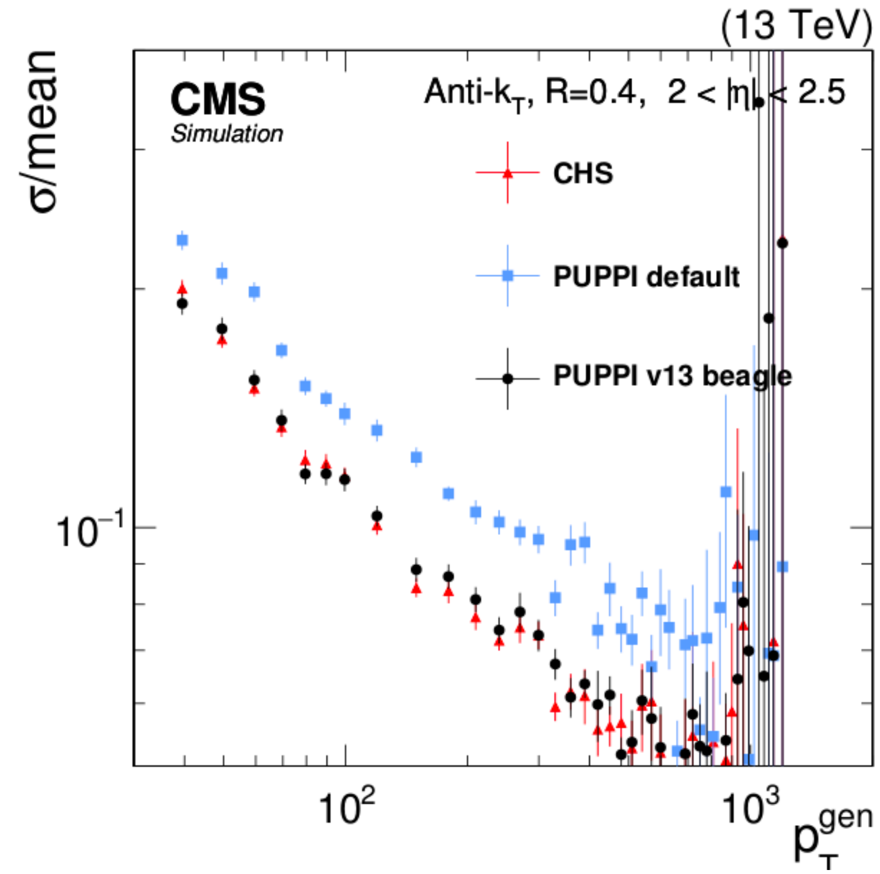
JER Ak4 Jets

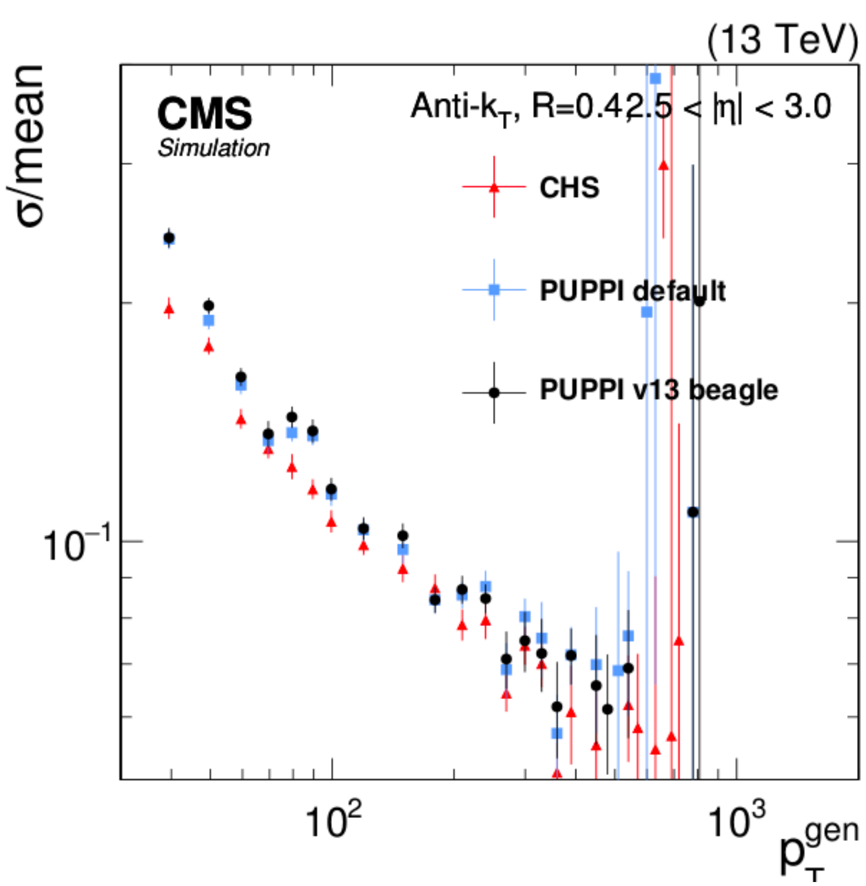
JER 2016 - Ak4

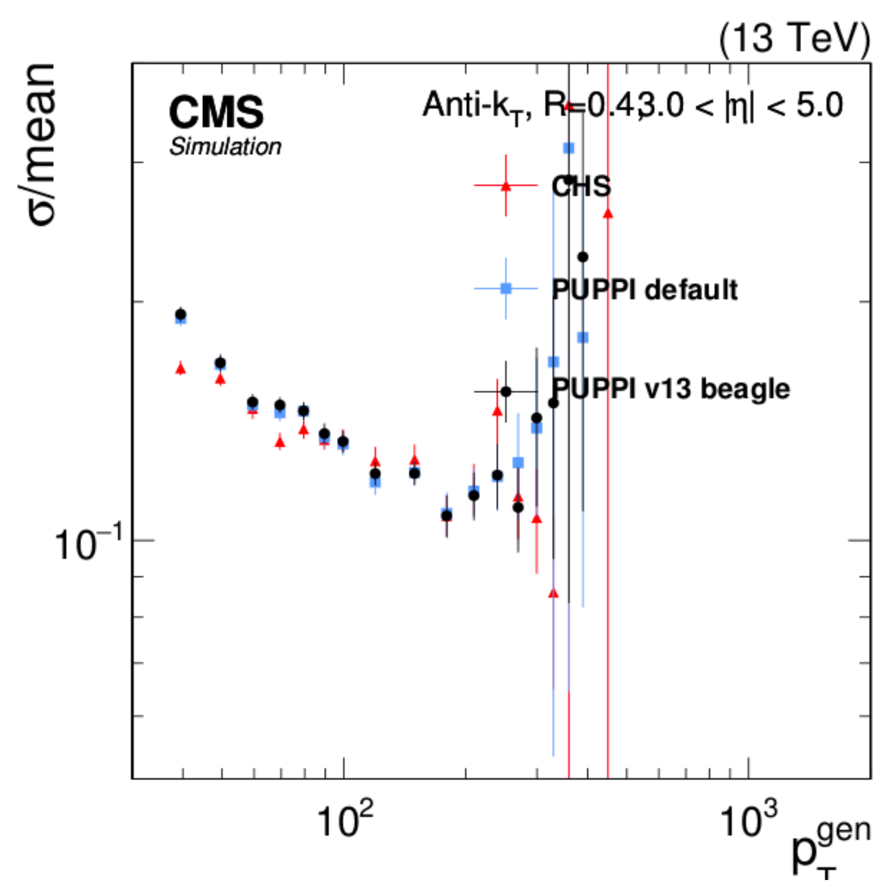
QCD sample







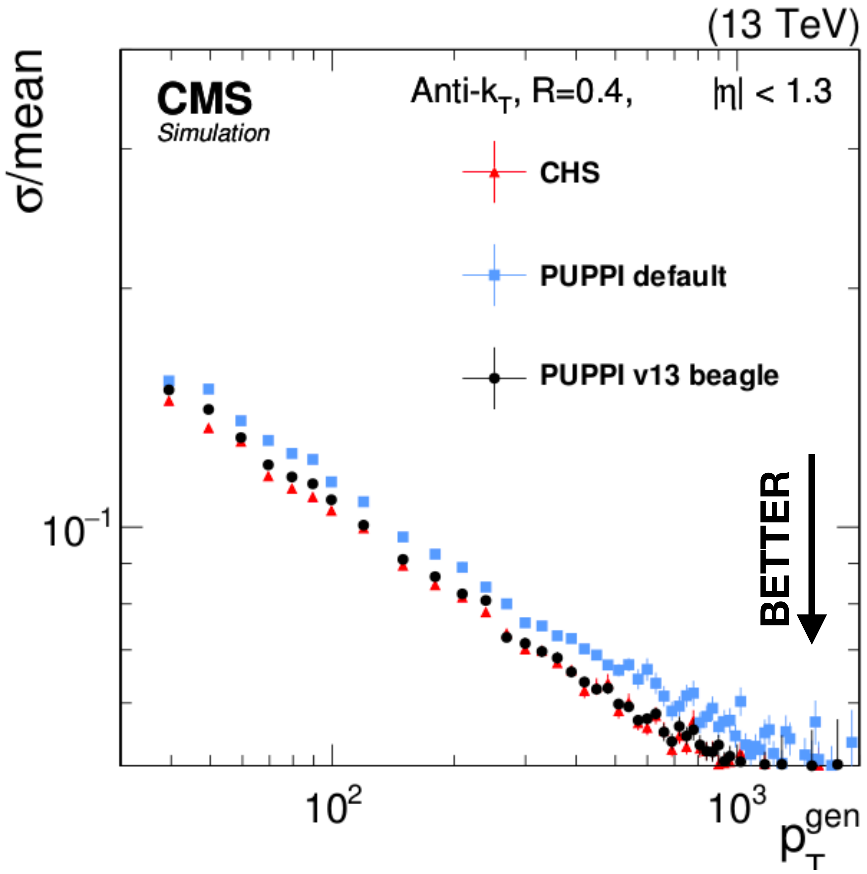


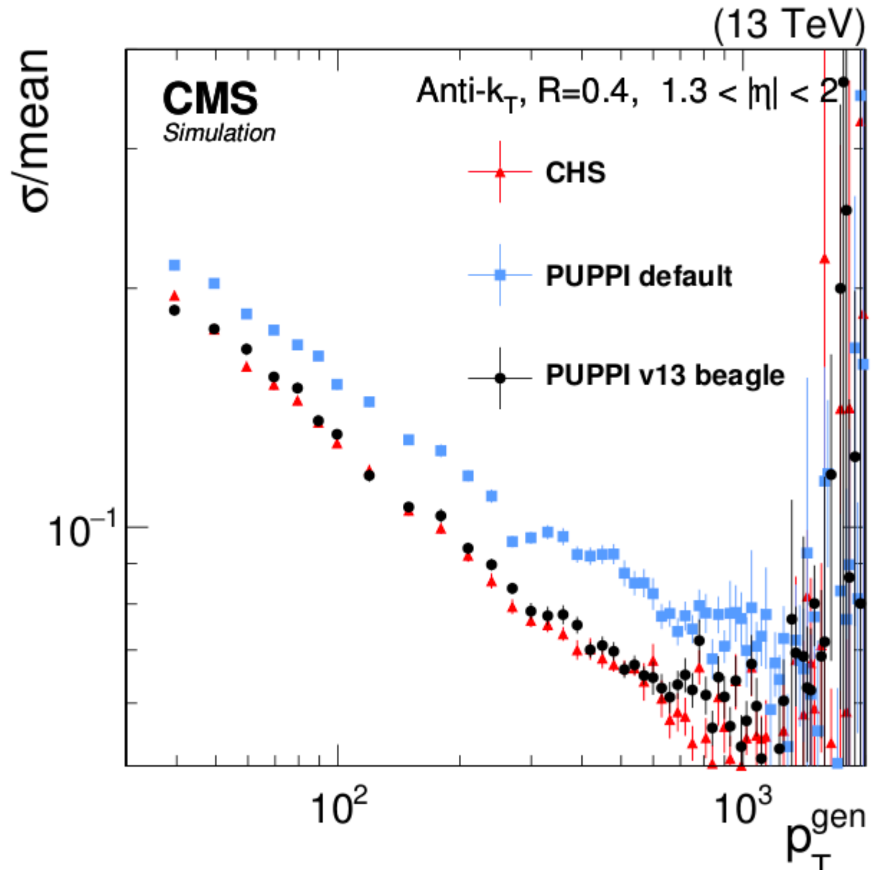


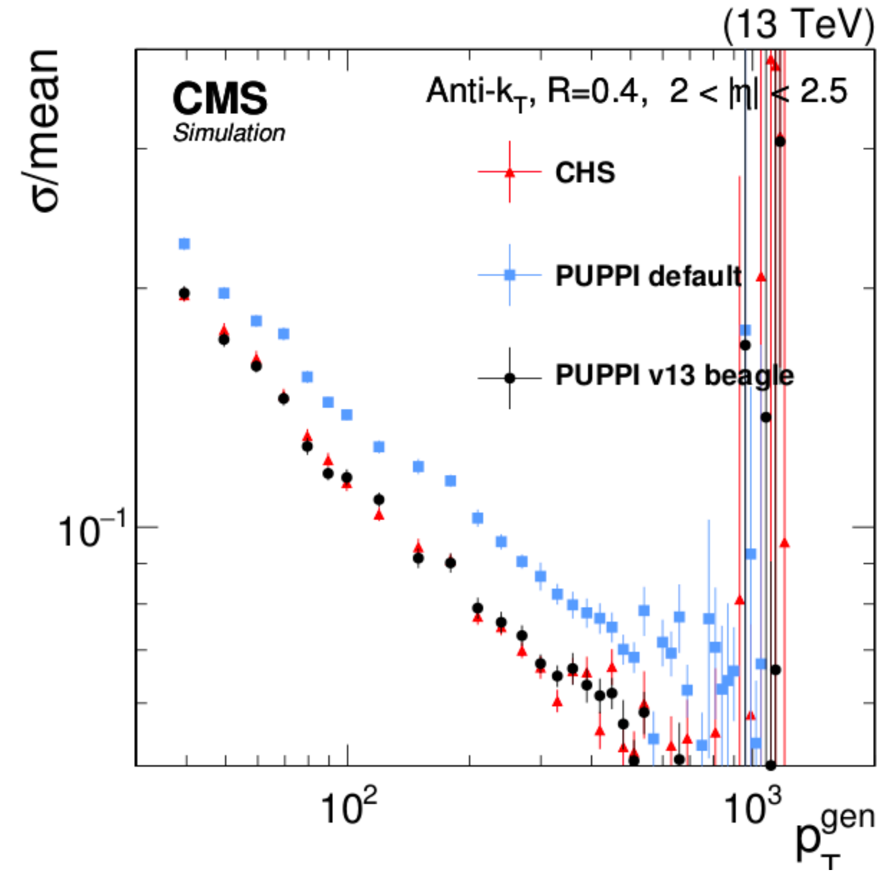
JER calculated with iterative gaussian fit

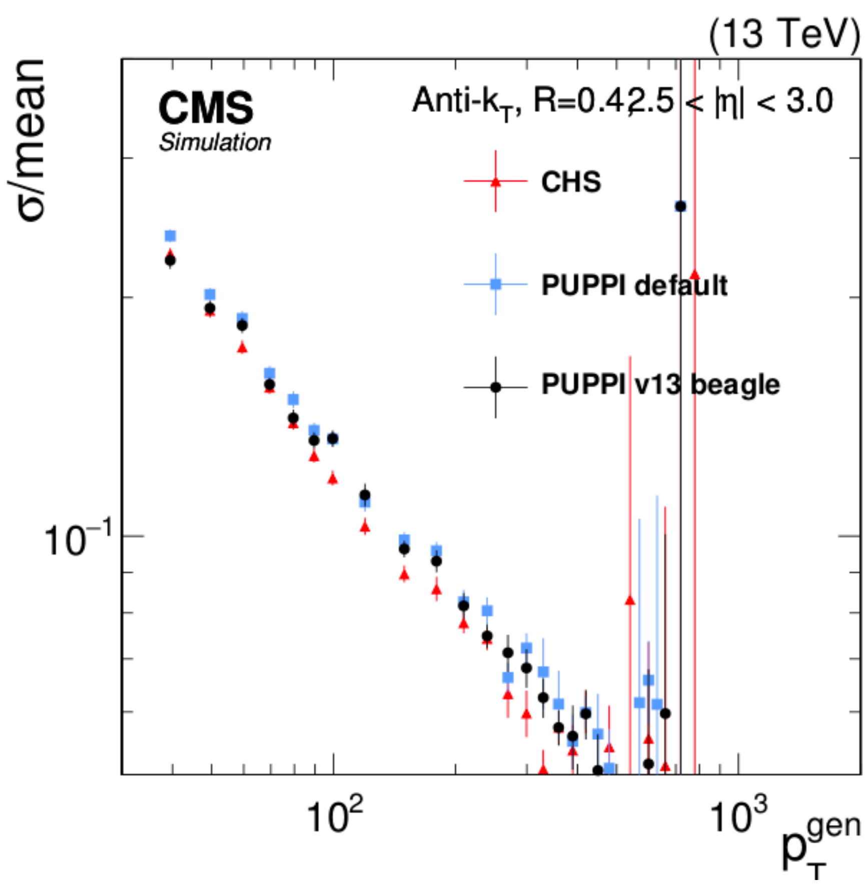
JER 2017 - Ak4

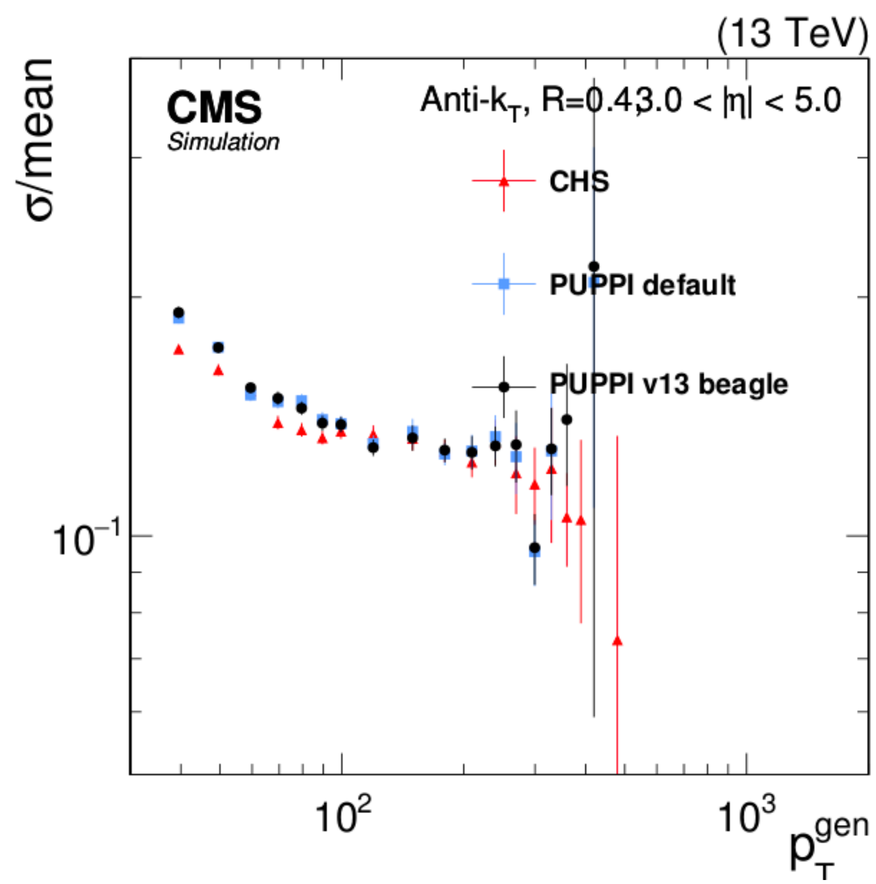
QCD sample







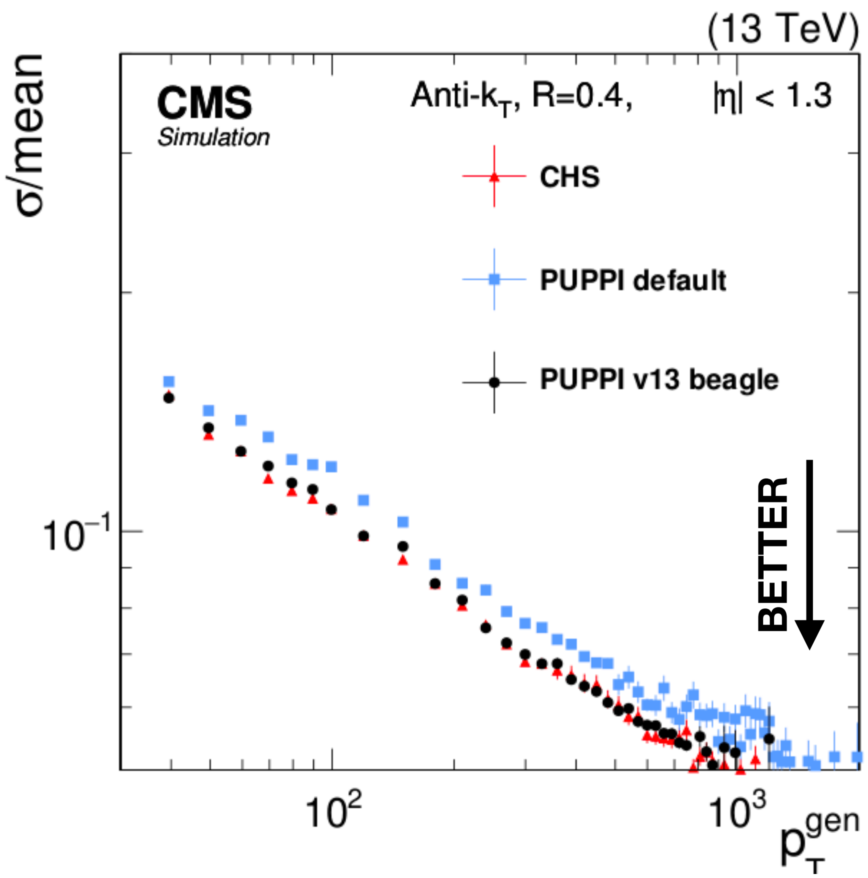


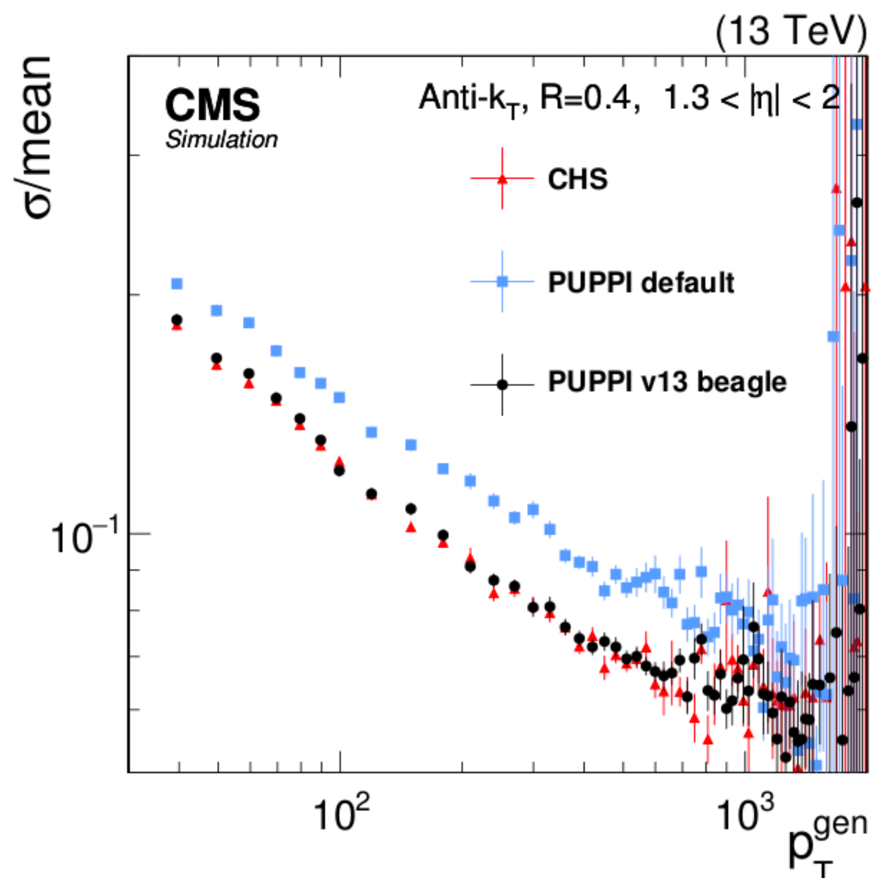


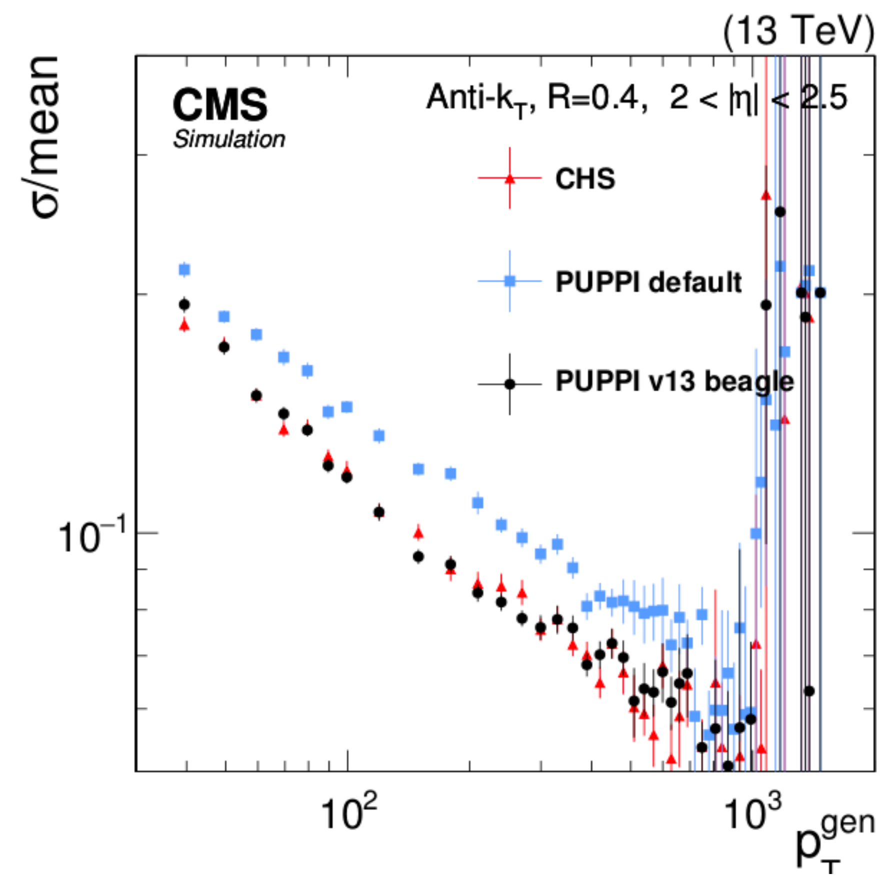
JER calculated with iterative gaussian fit

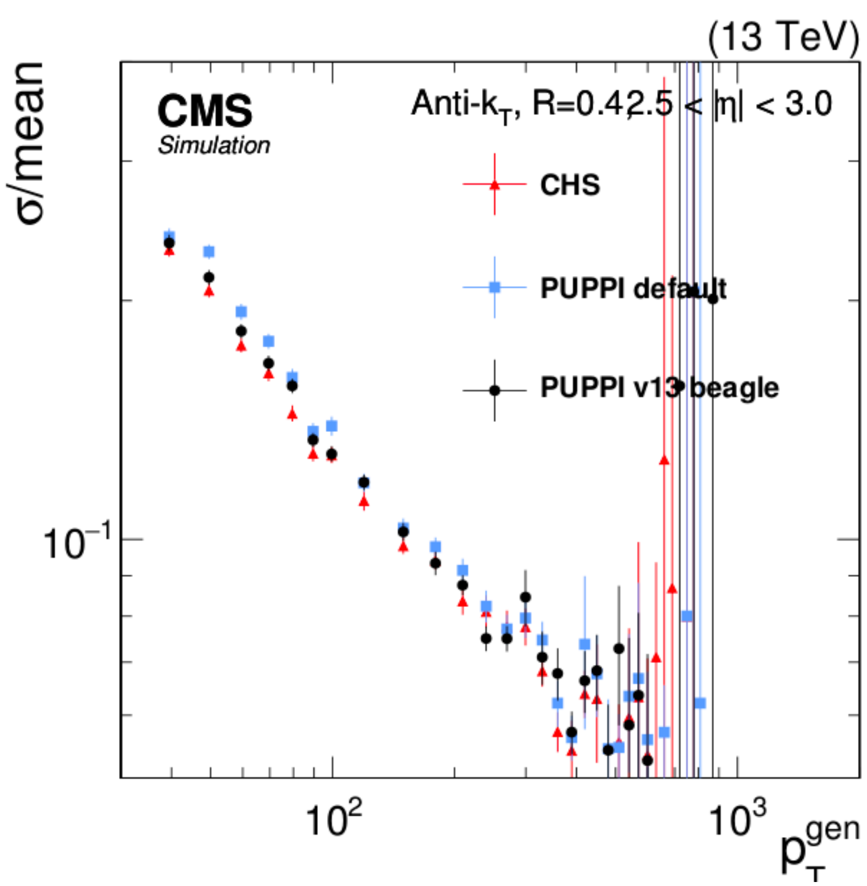
JER 2018 - Ak4

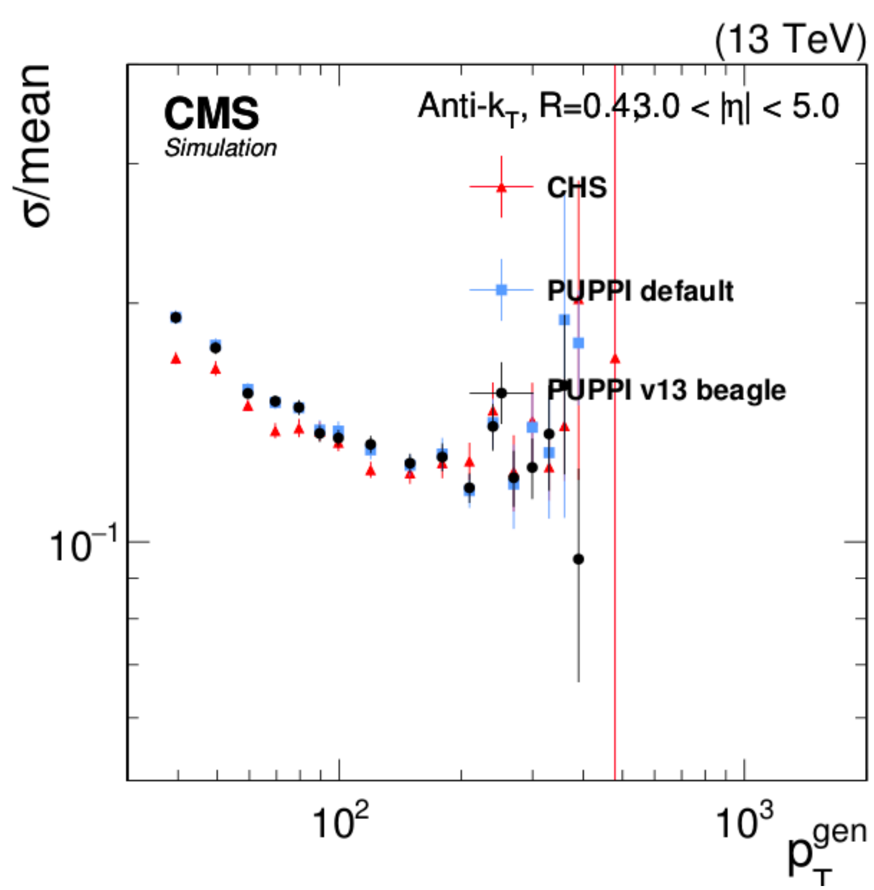
QCD sample







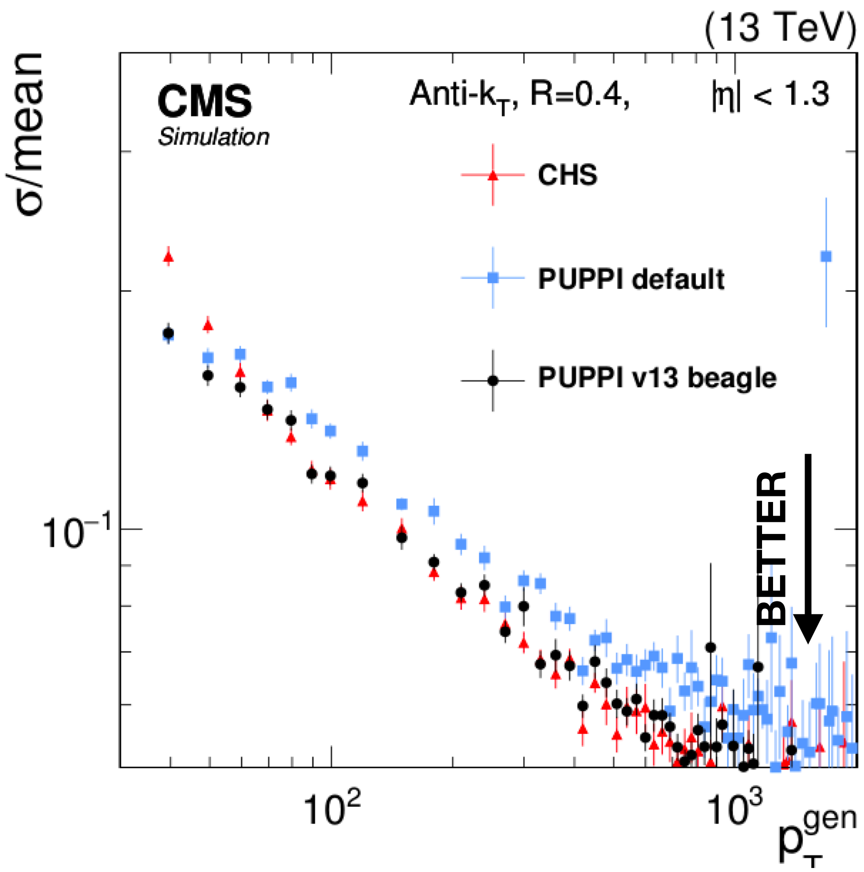


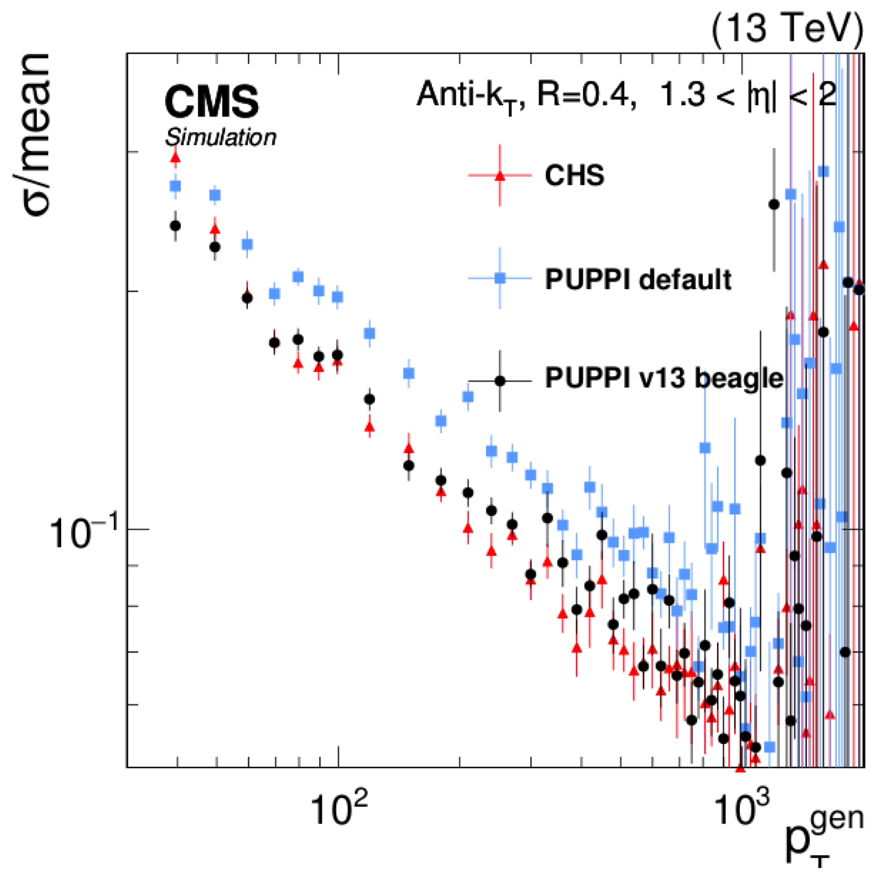


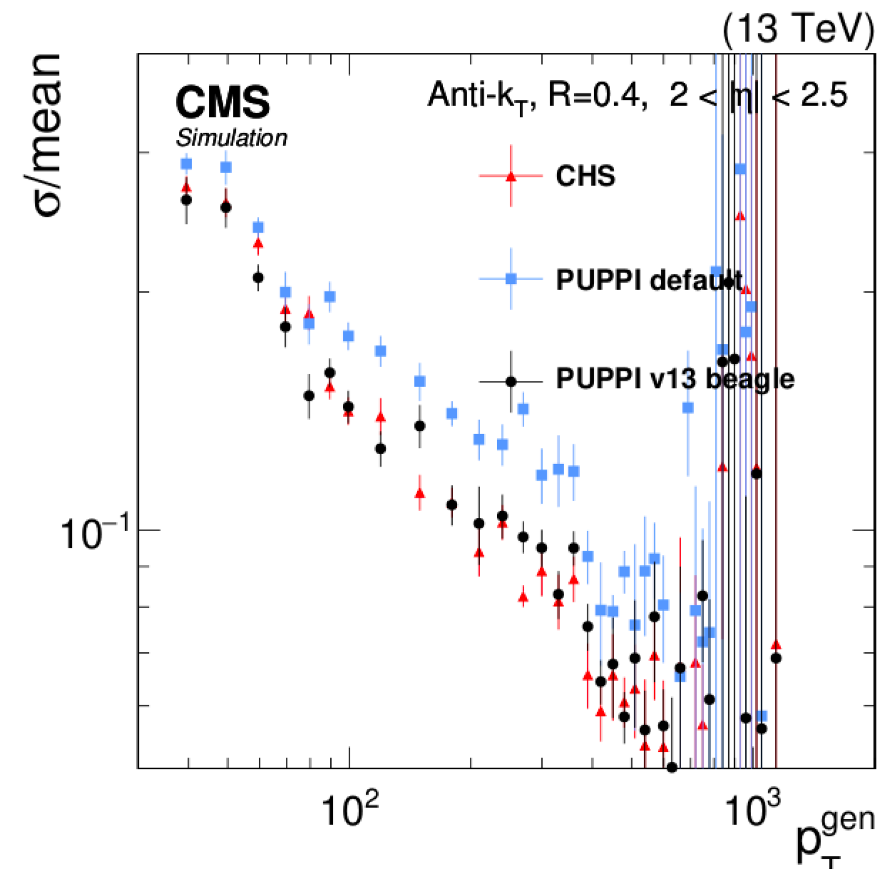
JER calculated with iterative gaussian fit

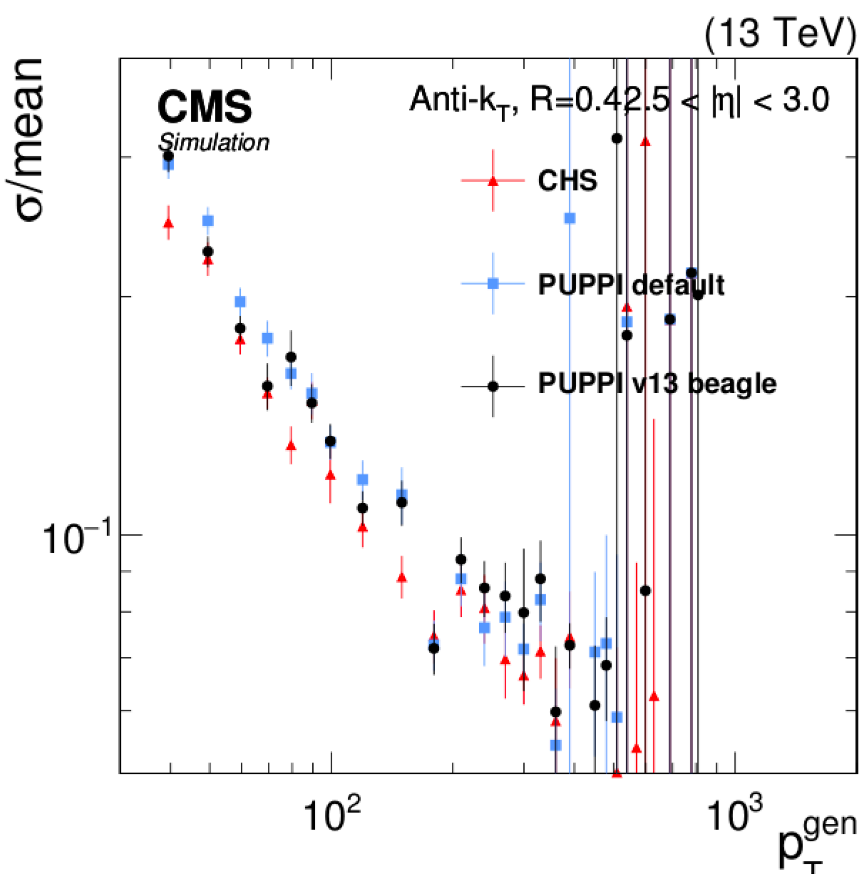
JER 2016 - Ak4 PU>50

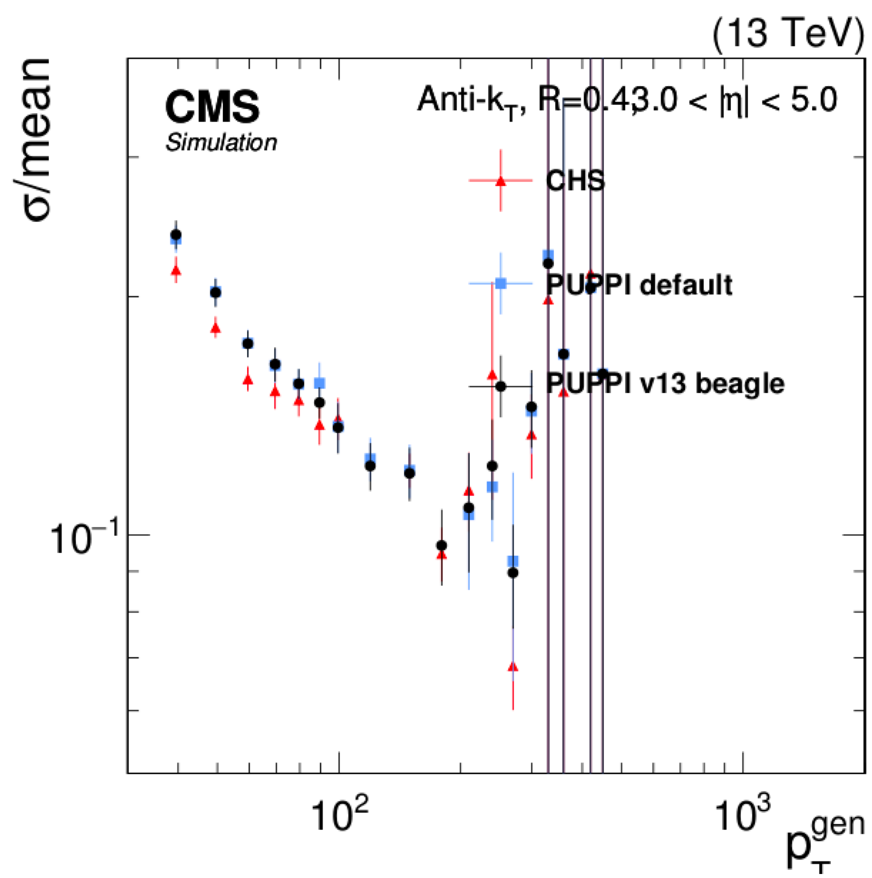
<u>QCD sample</u>







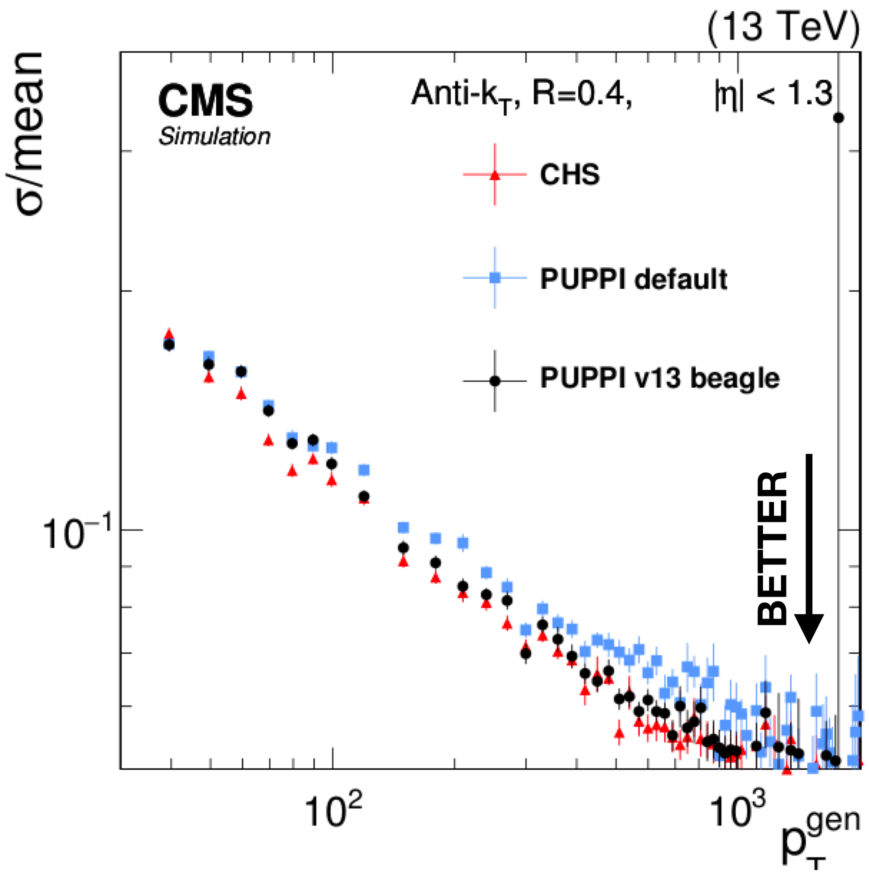


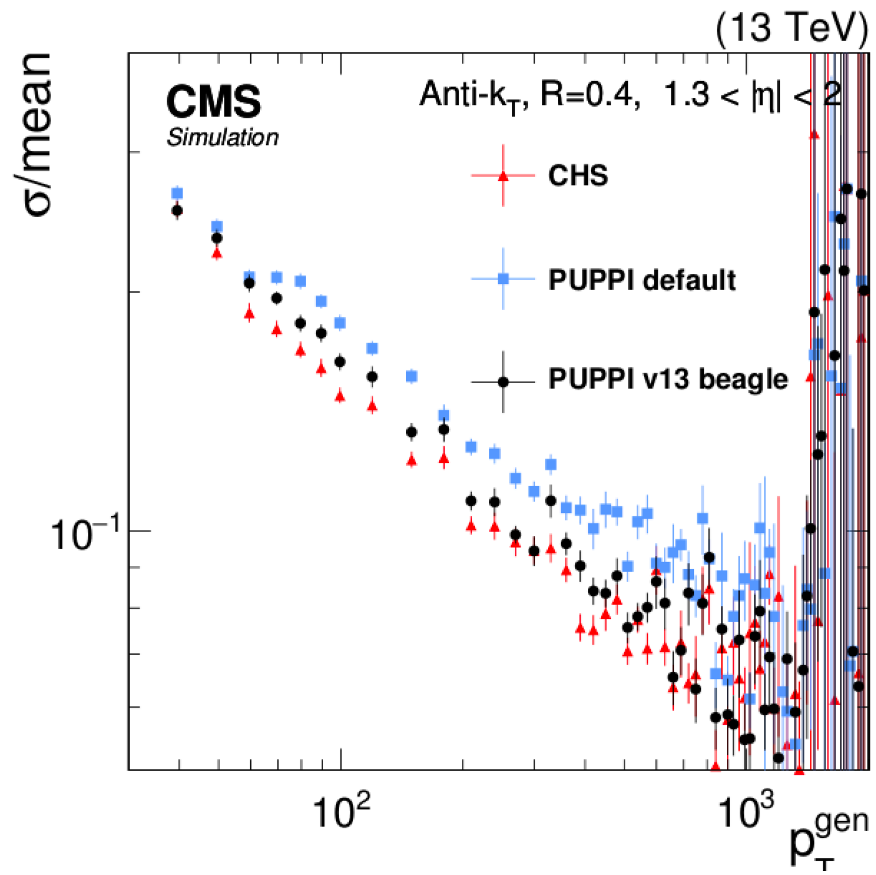


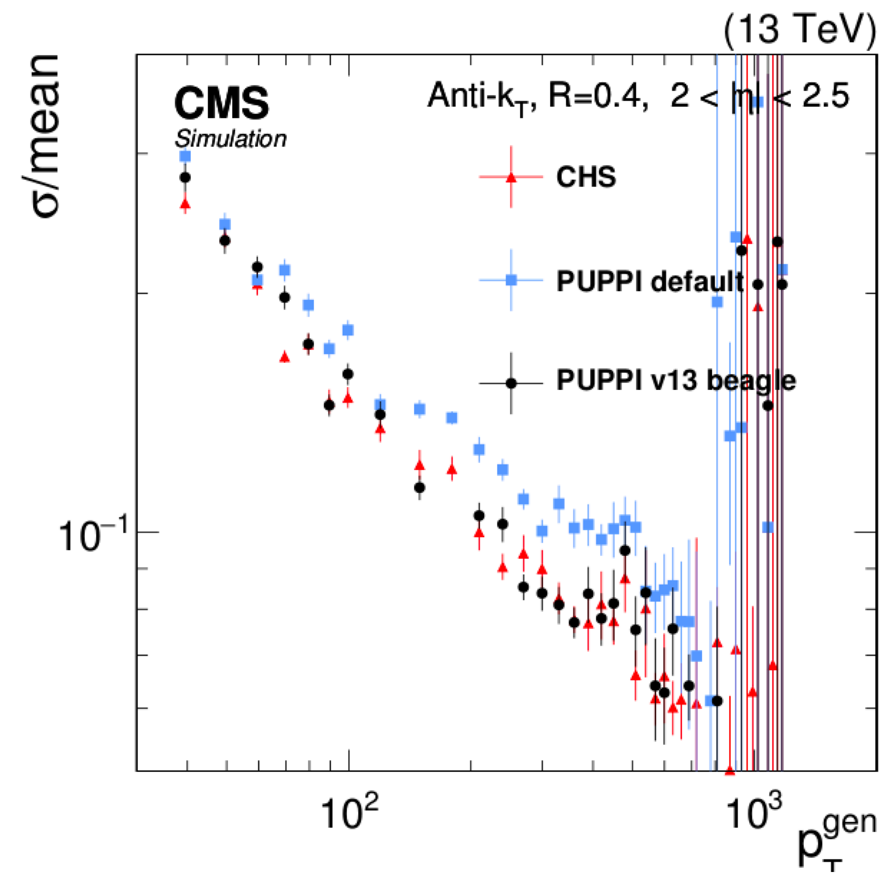
JER calculated with iterative gaussian fit

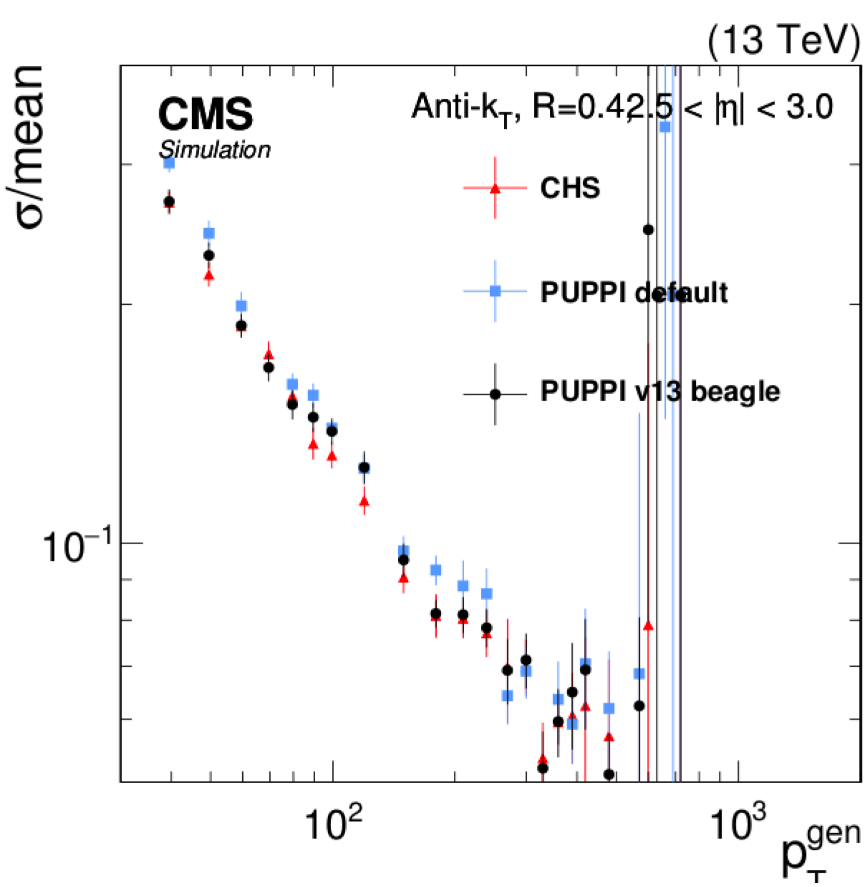
JER 2017 - Ak4 PU>50

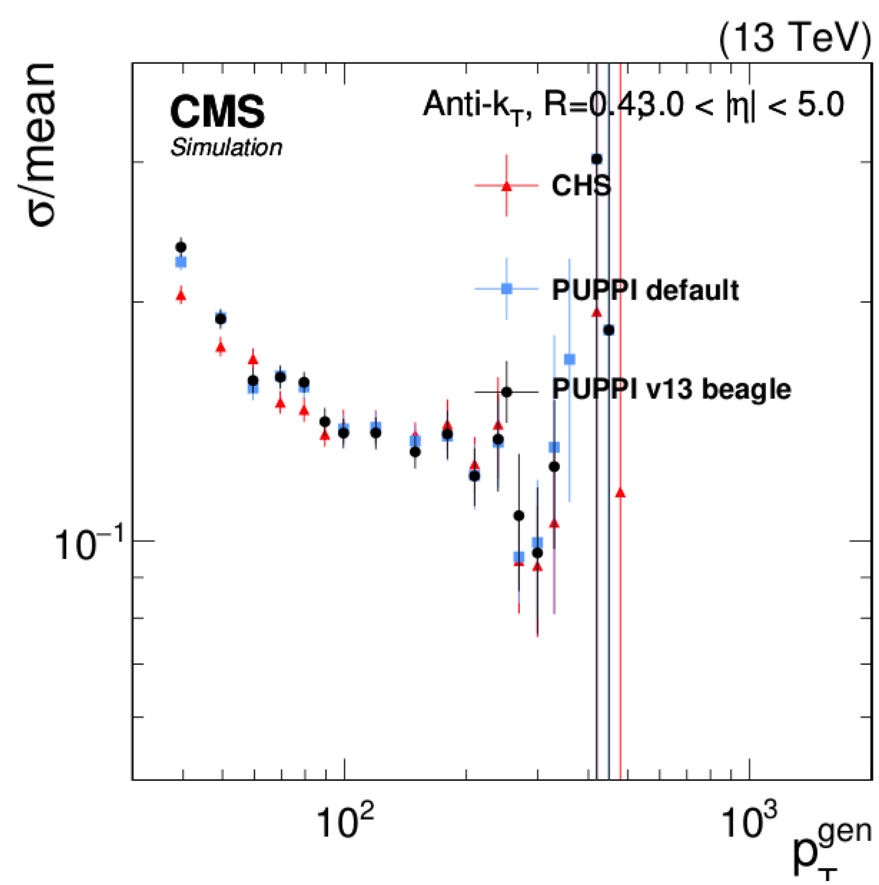
QCD sample







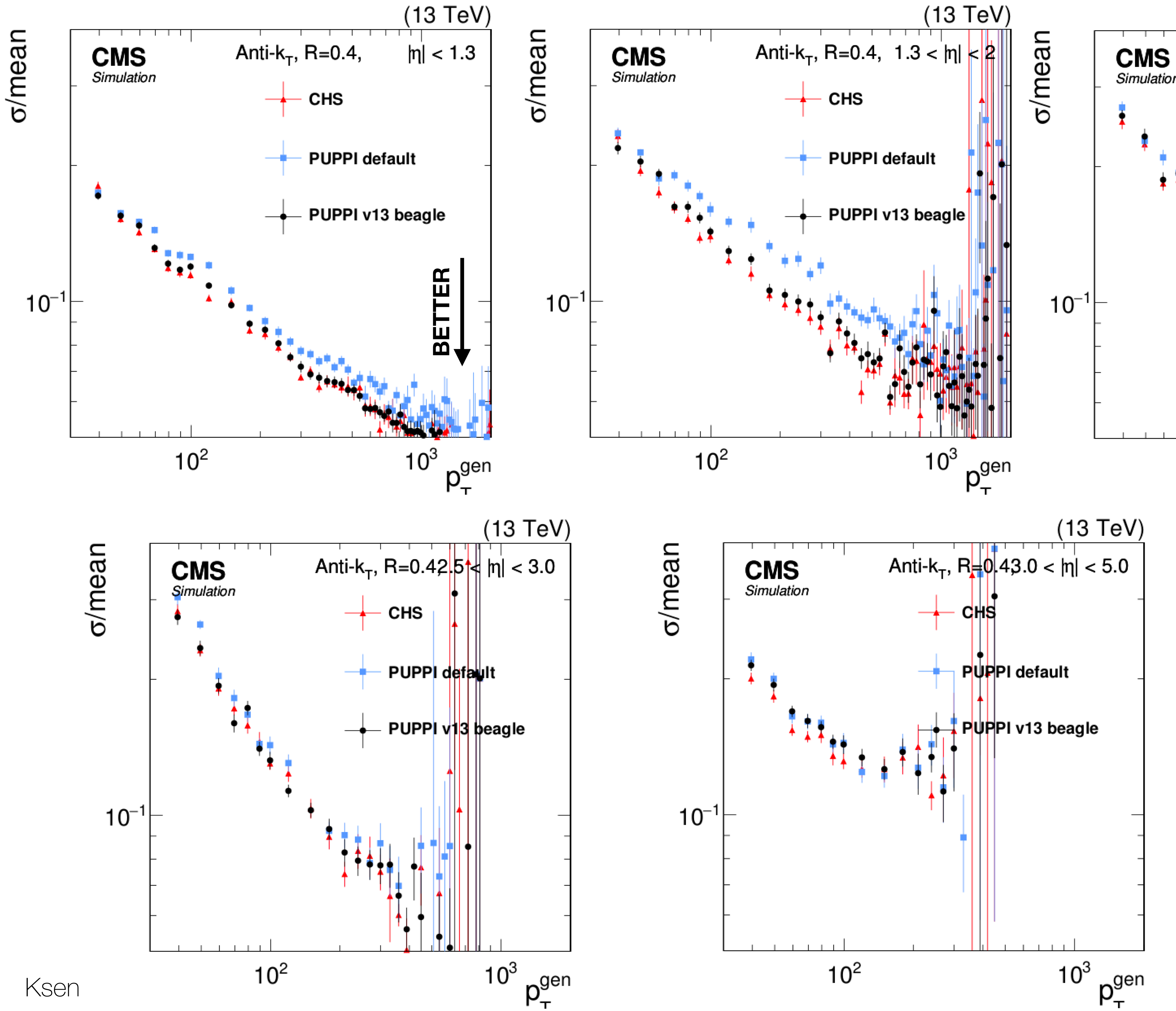




JER calculated with iterative gaussian fit

JER 2018 - Ak4 PU>50

QCD sample

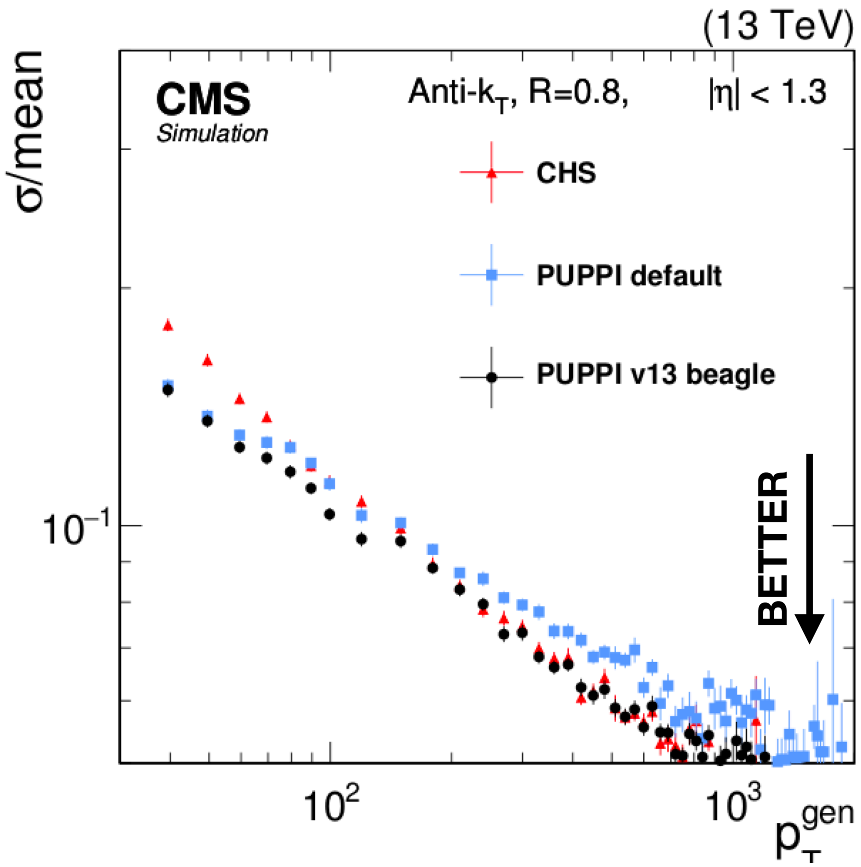


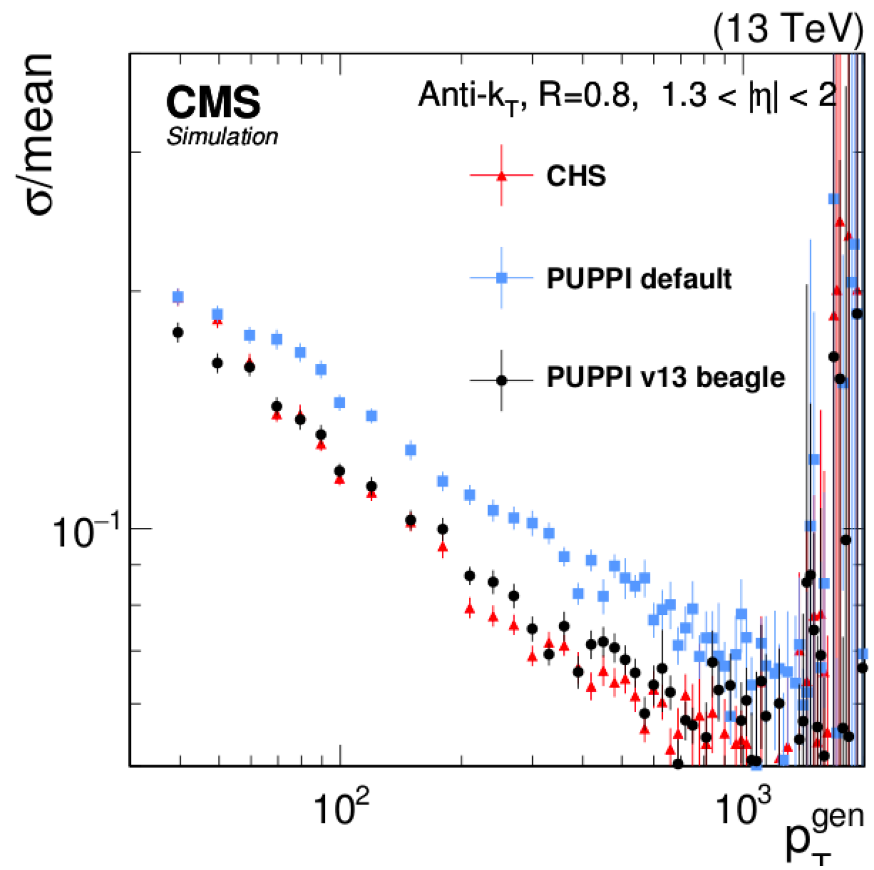
JER calculated with iterative gaussian fit

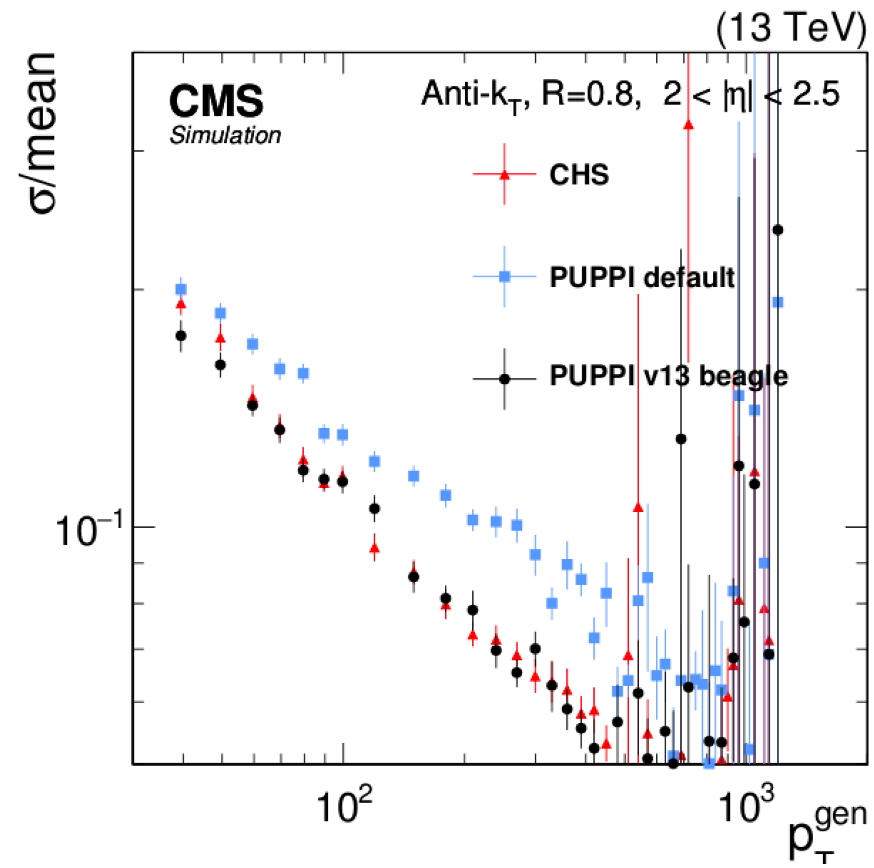
JER Ak8 Jets

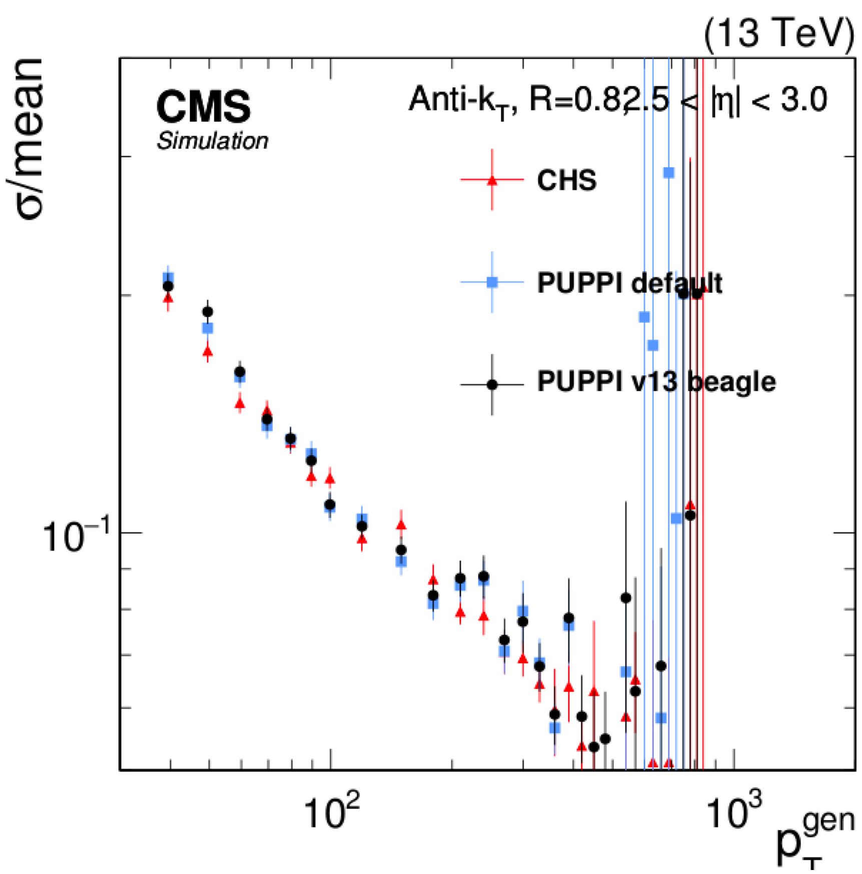
JER 2016 - Ak8

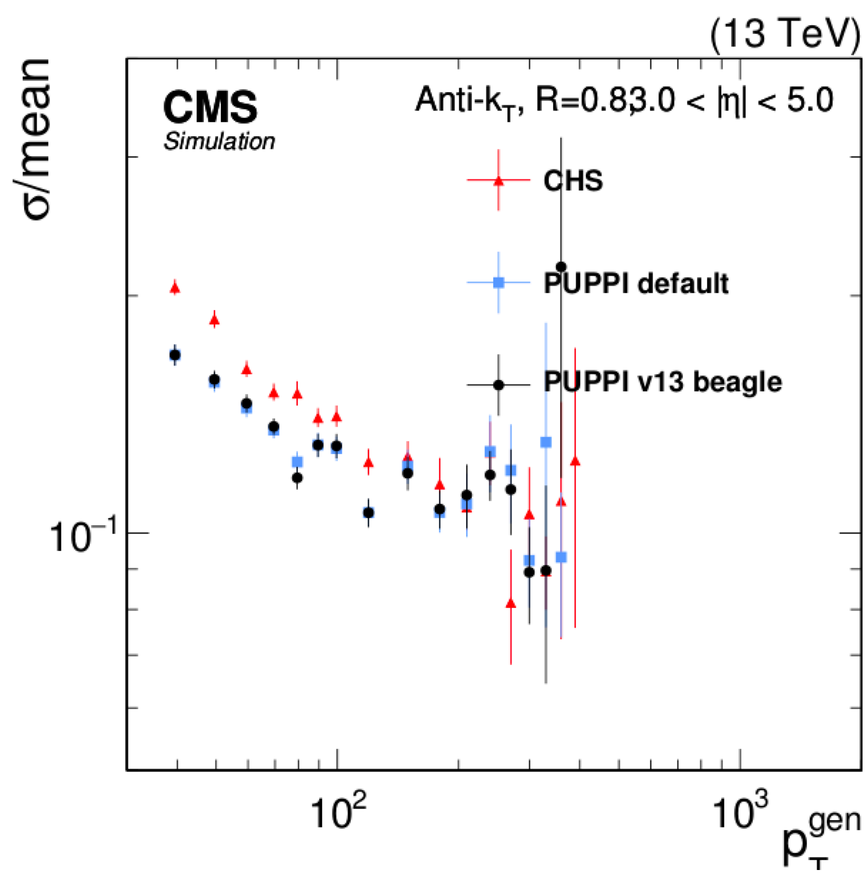
<u>QCD sample</u>







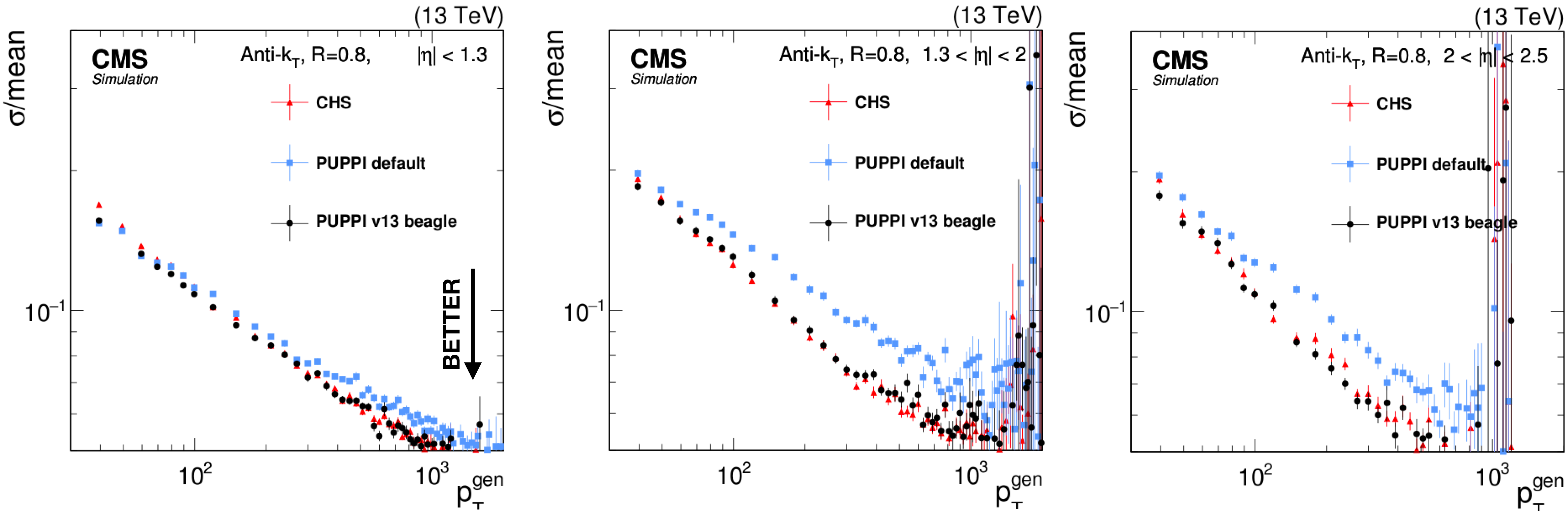


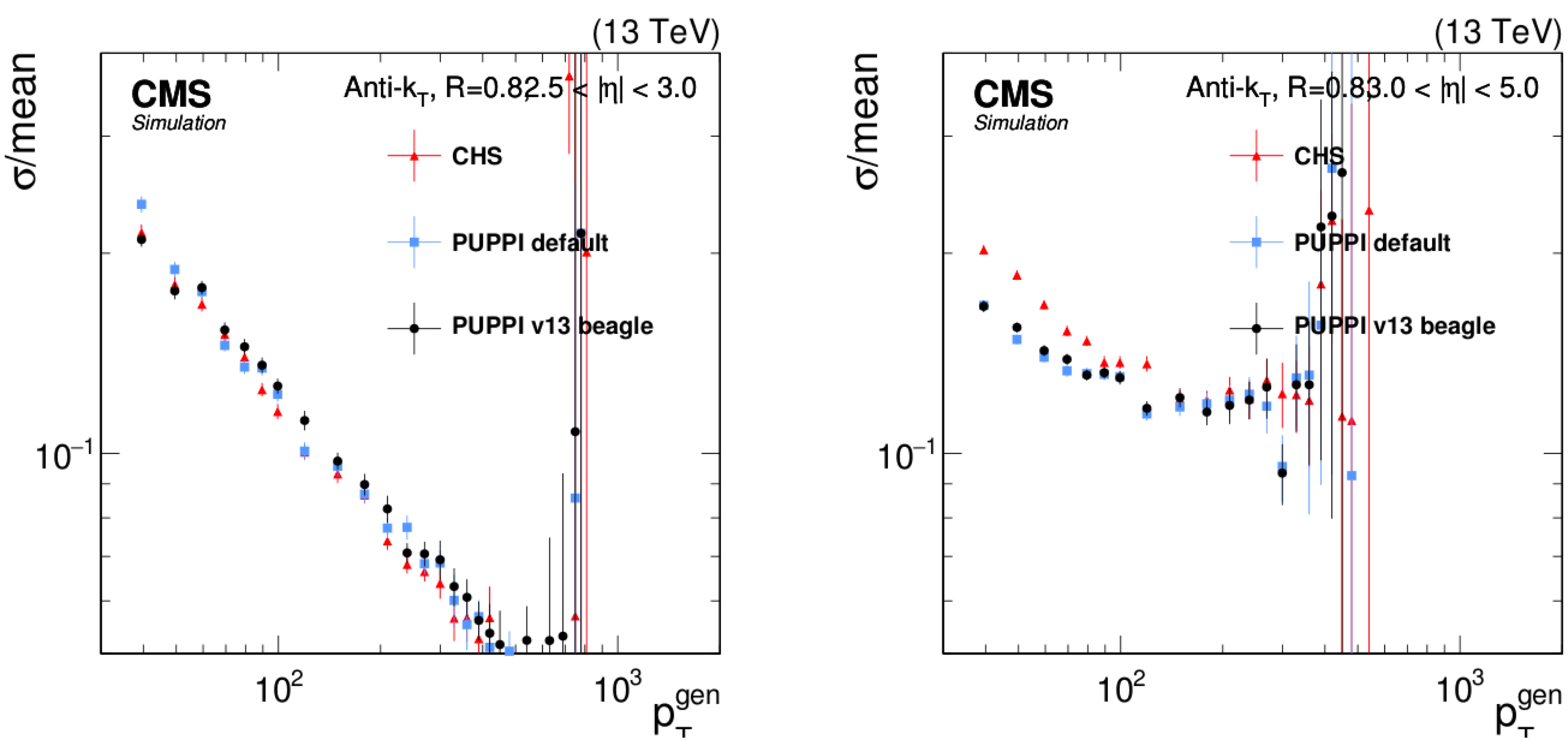


JER calculated with iterative gaussian fit

JER 2017 - Ak8

<u>QCD sample</u>

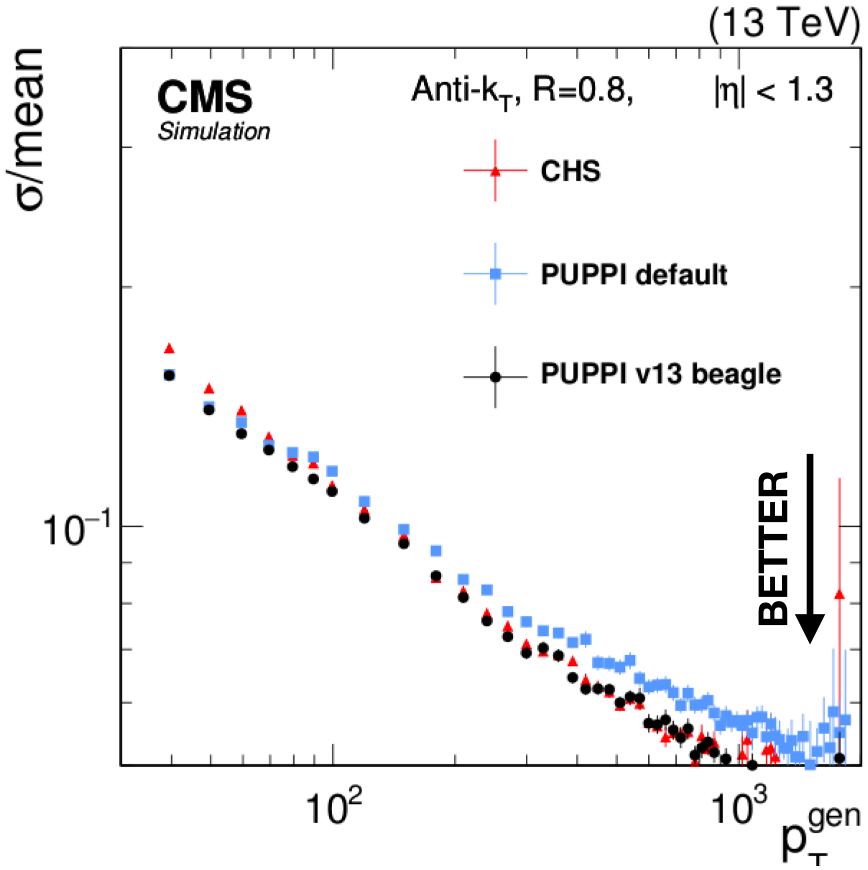


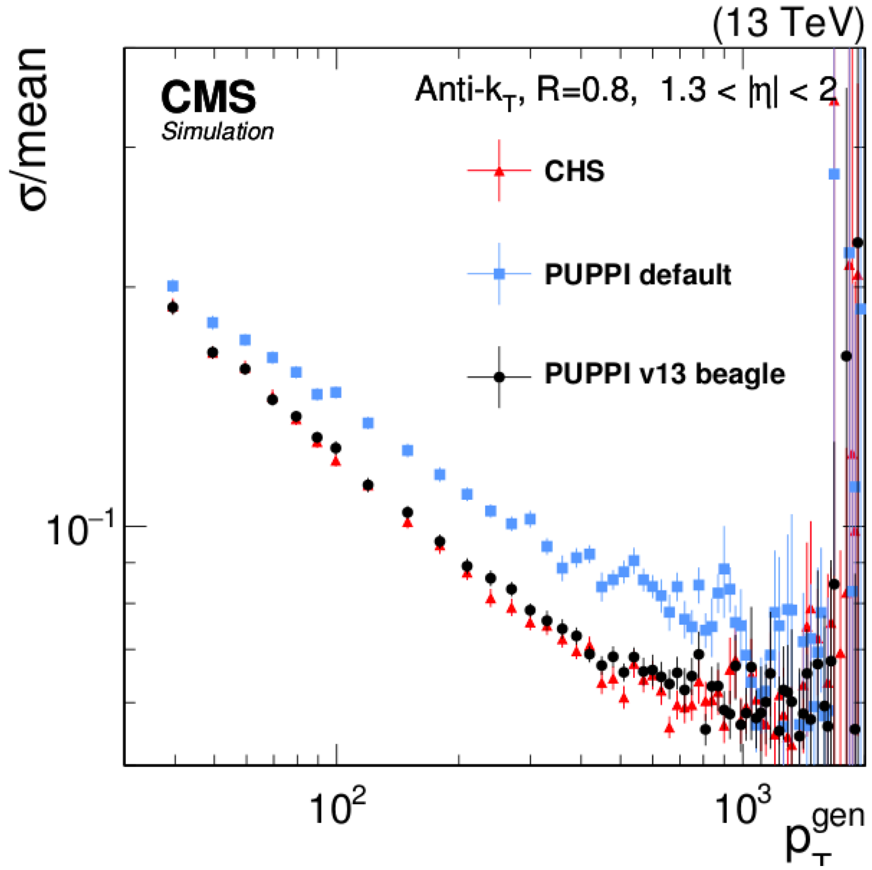


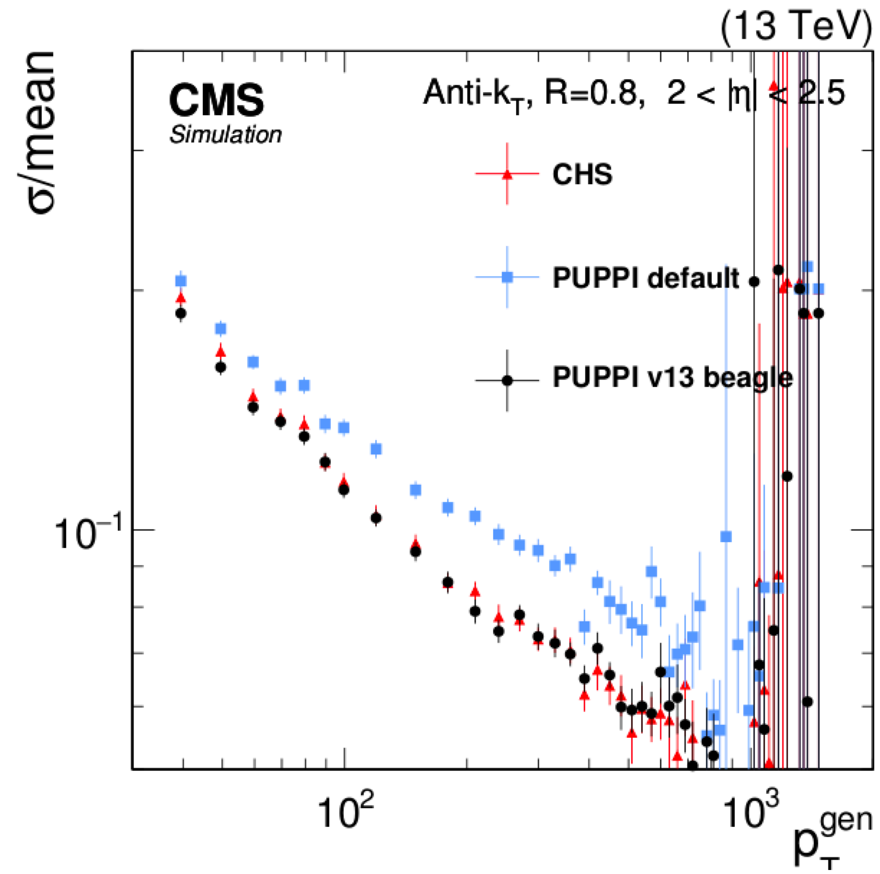
JER calculated with iterative gaussian fit

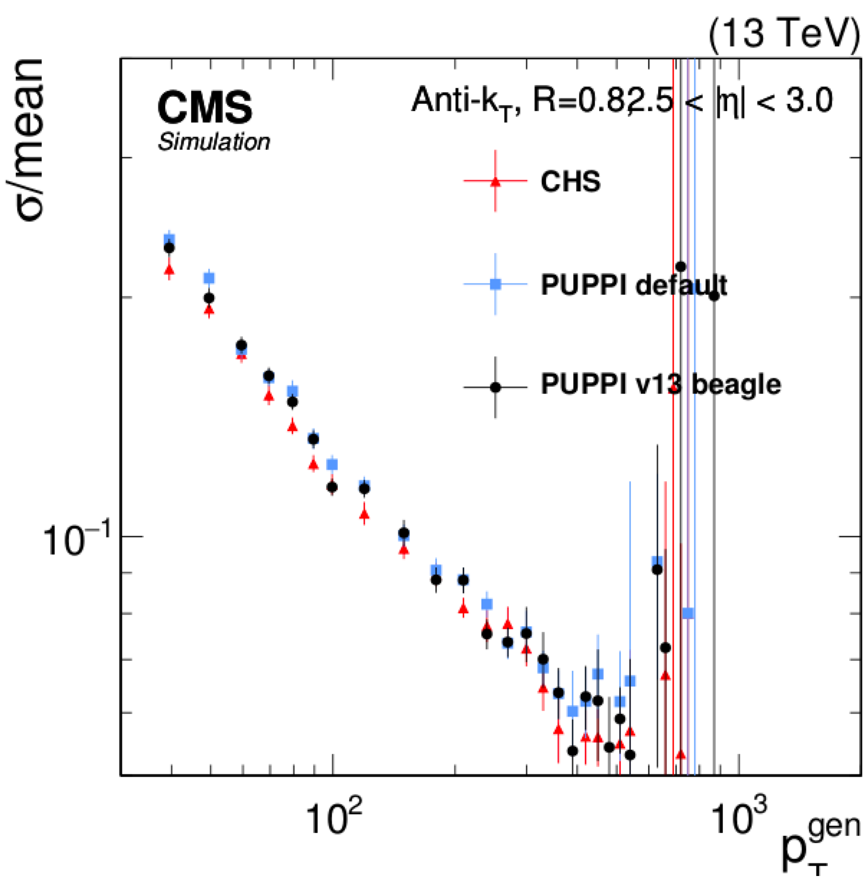
JER 2018 - Ak8

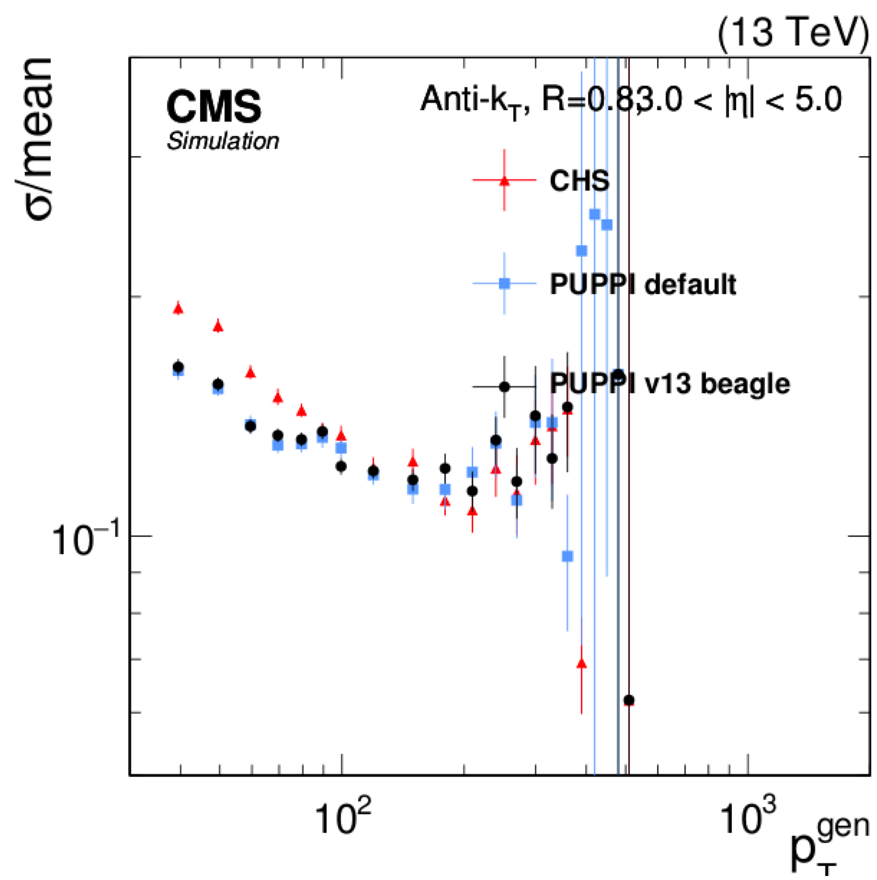
QCD sample







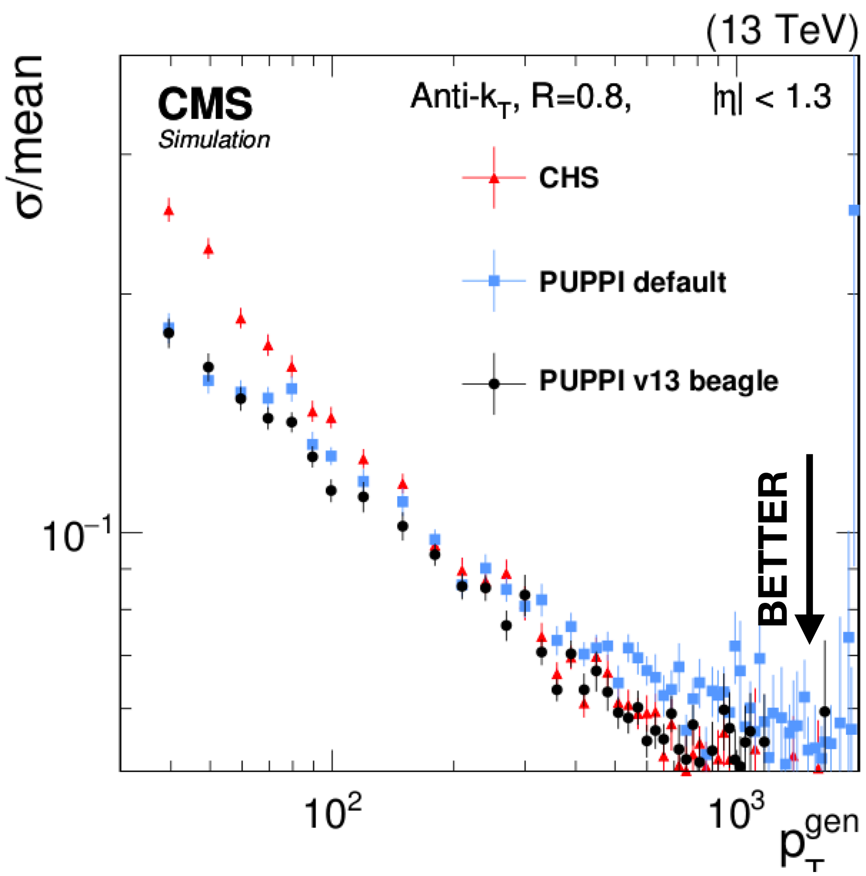


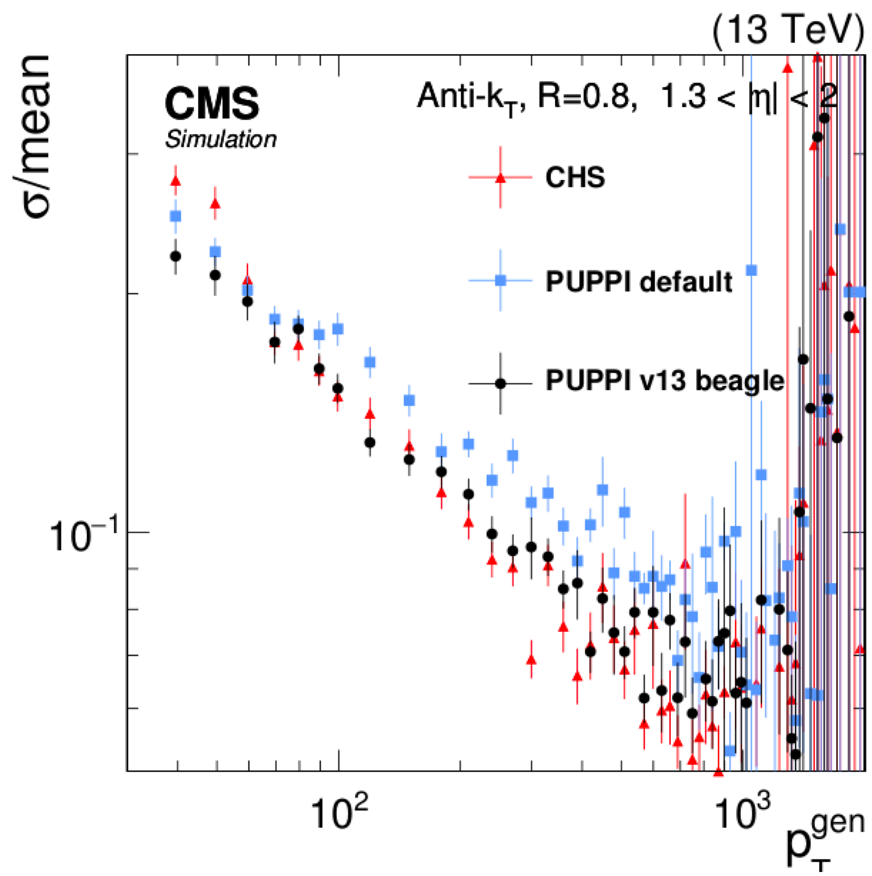


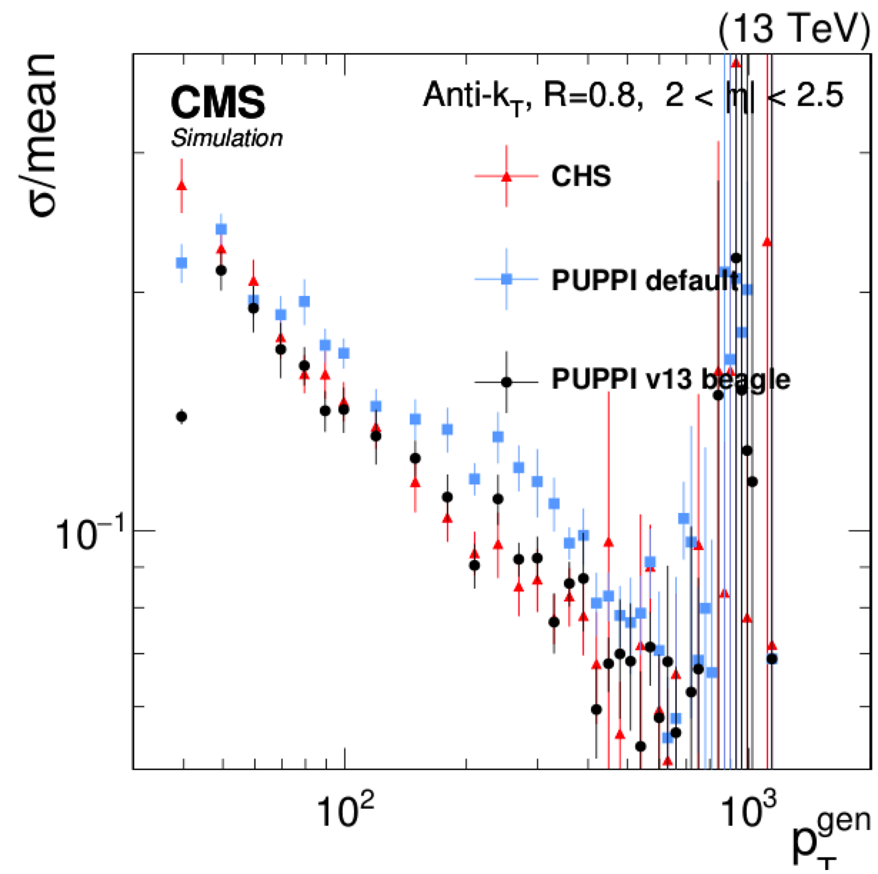
JER calculated with iterative gaussian fit

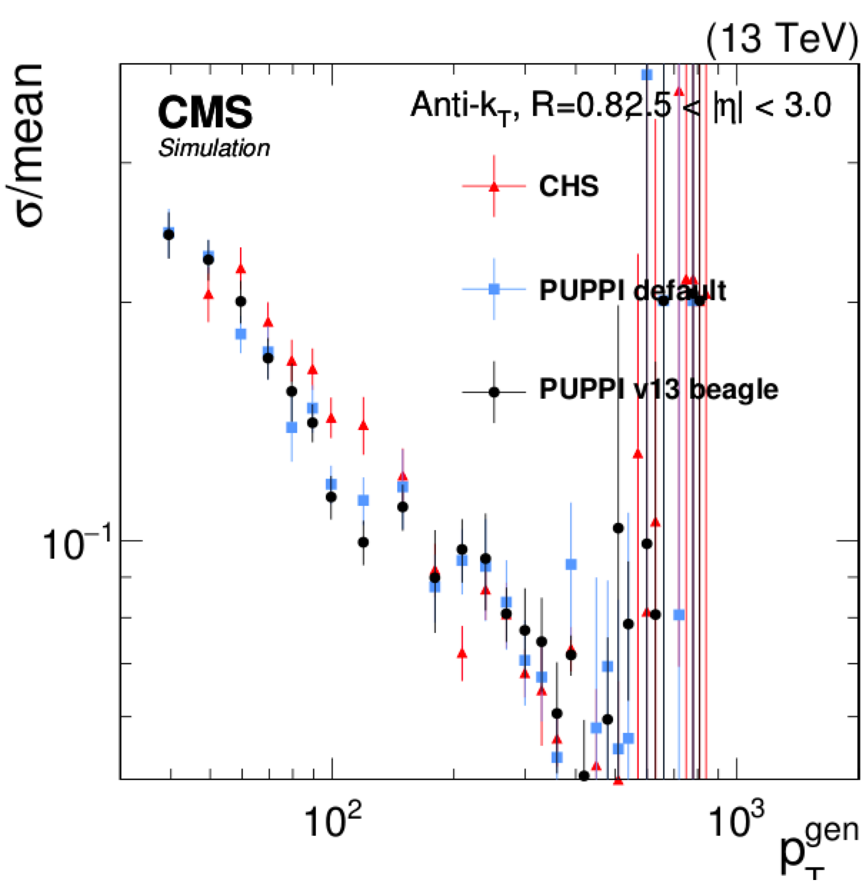
JER 2016 - Ak8 PU>50

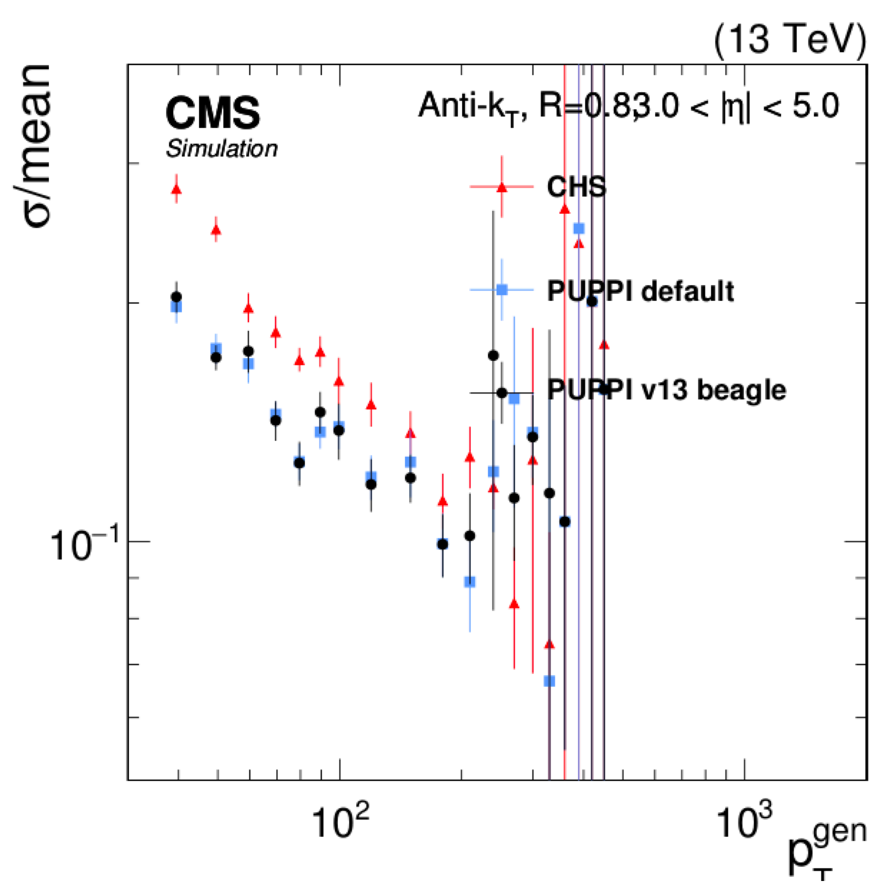
QCD sample







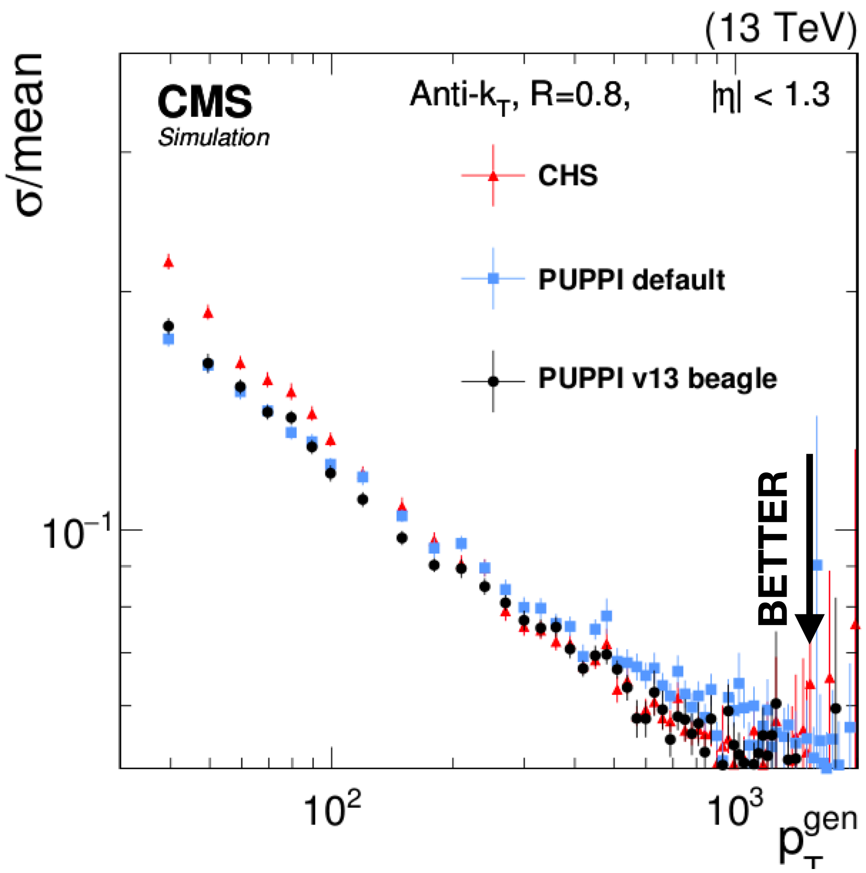


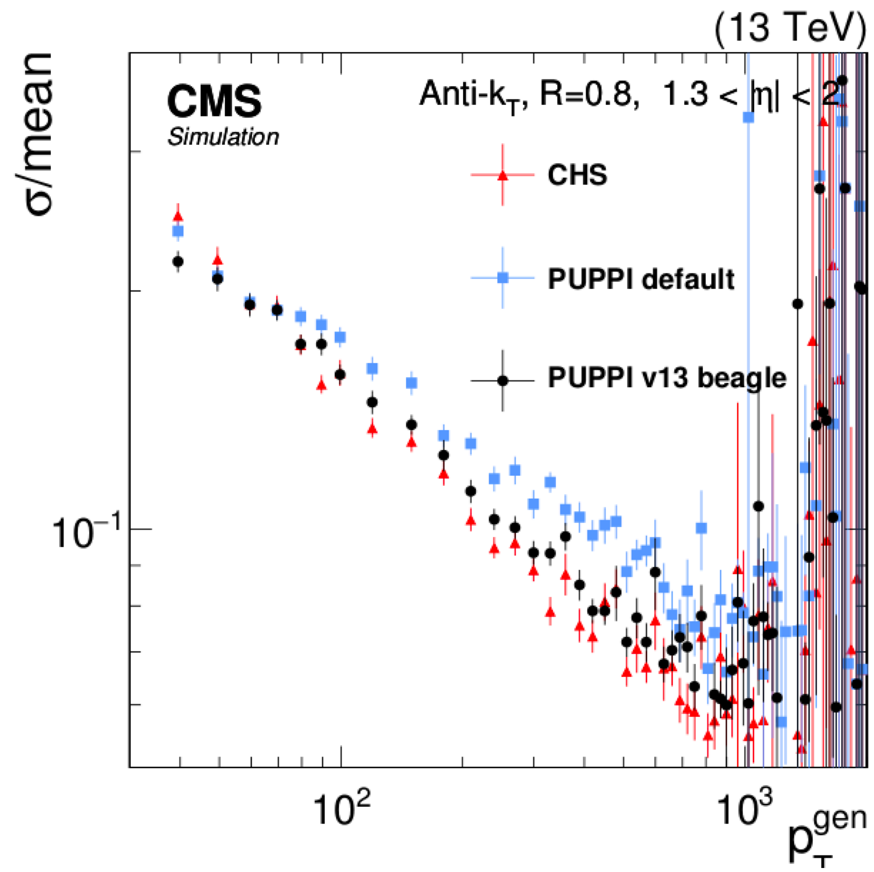


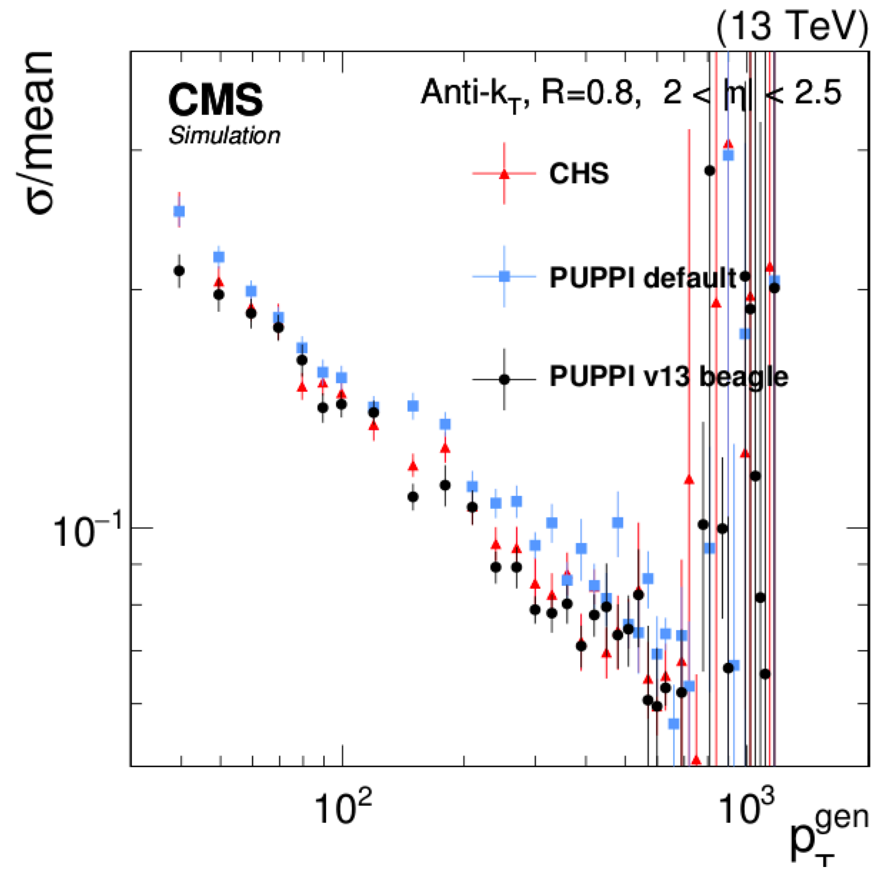
JER calculated with iterative gaussian fit

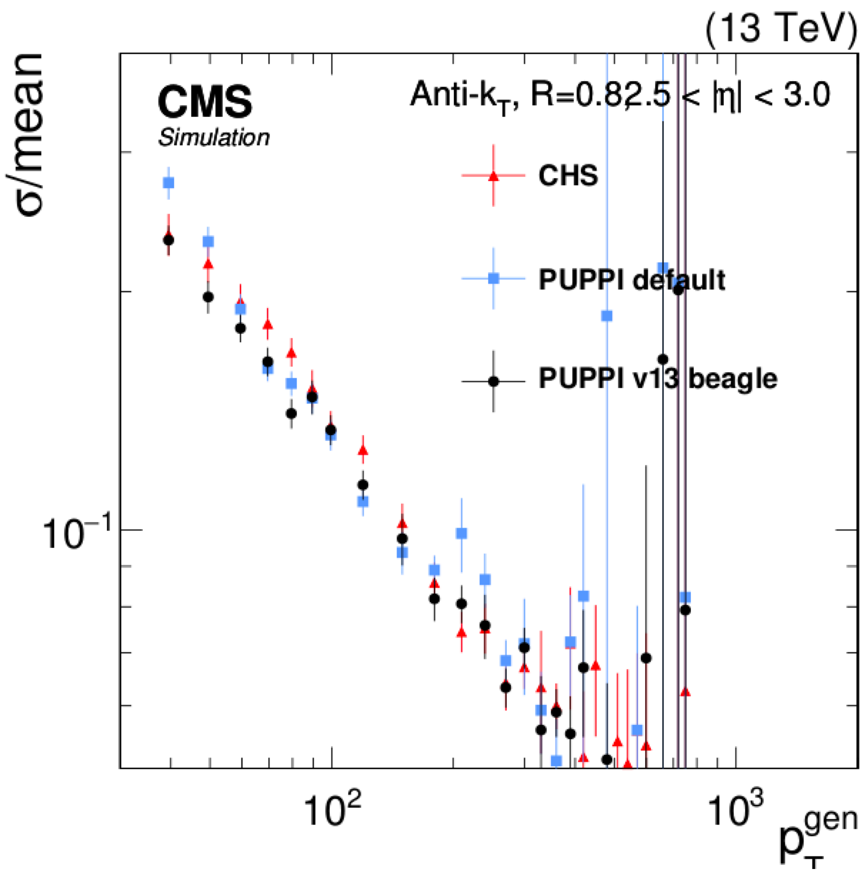
JER 2017 - Ak8 PU>50

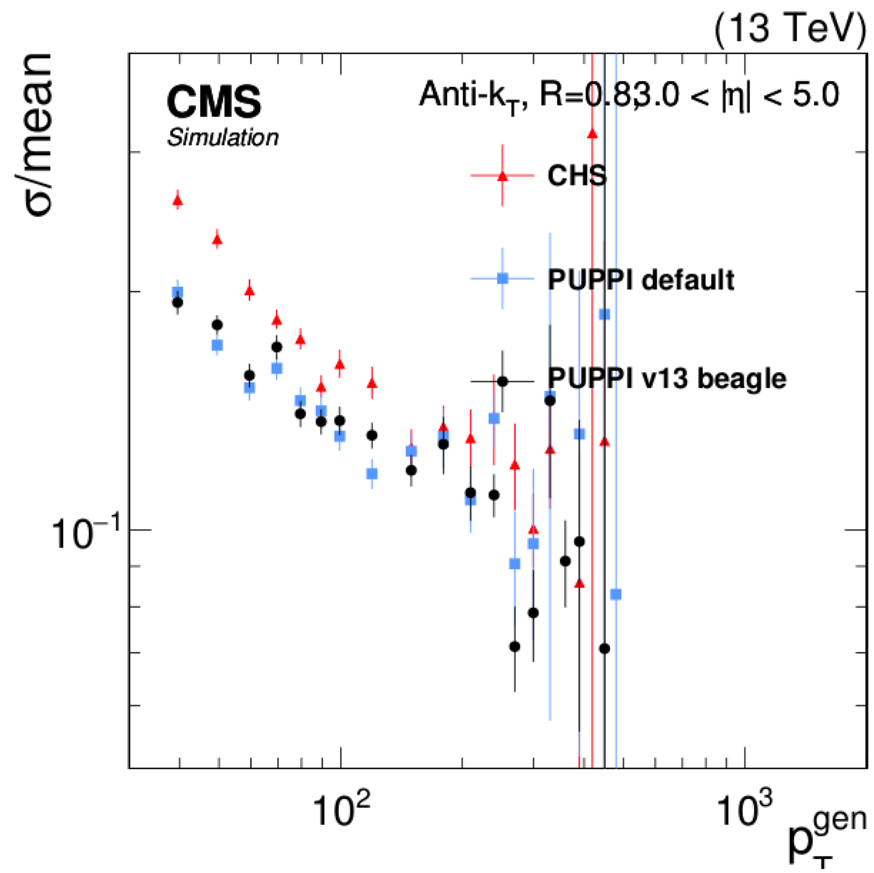
QCD sample







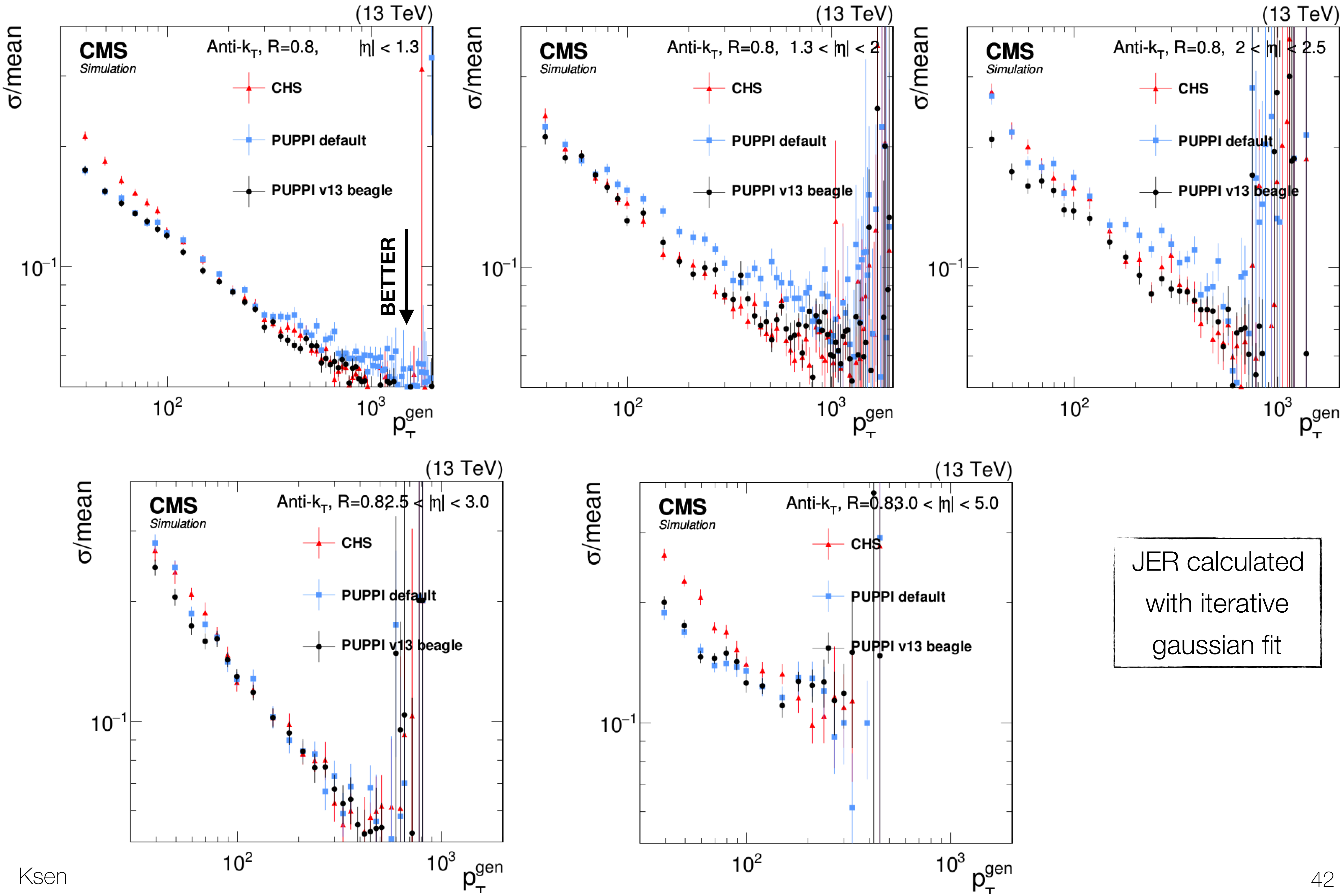




JER calculated with iterative gaussian fit

JER 2018 - Ak8 PU>50

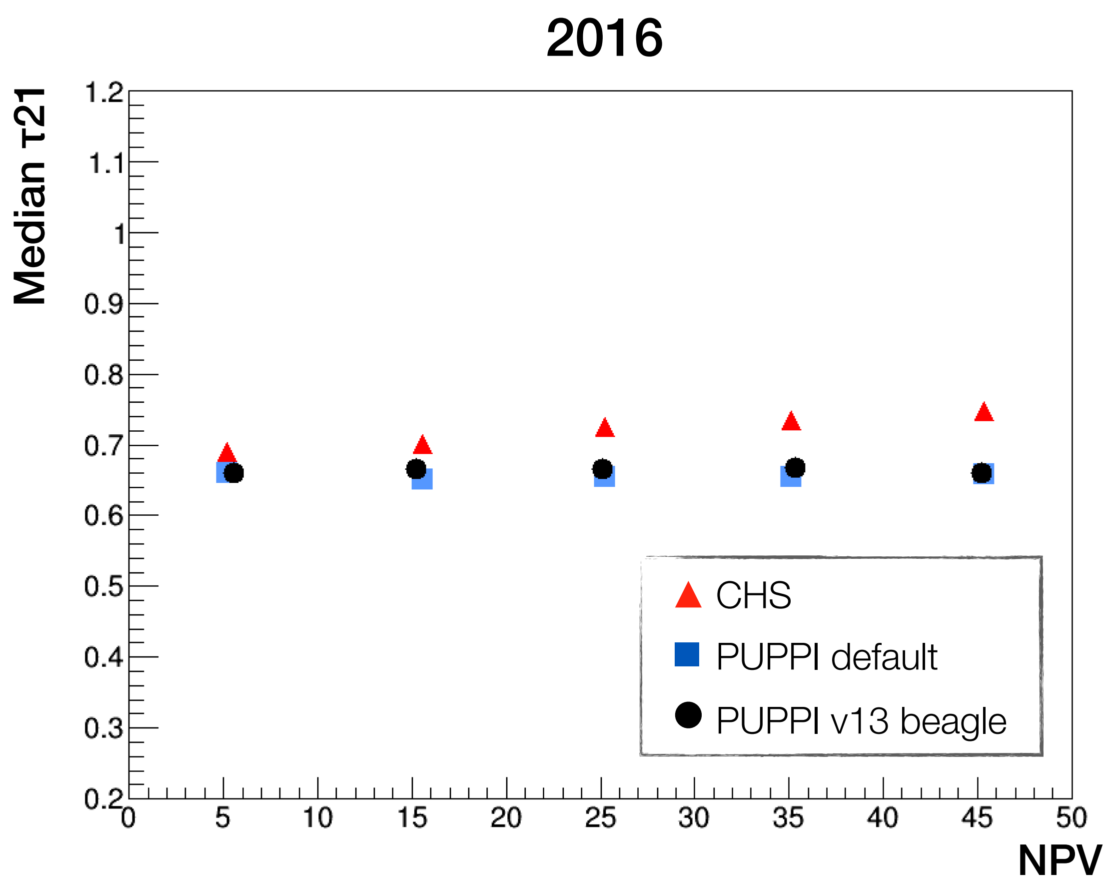
QCD sample

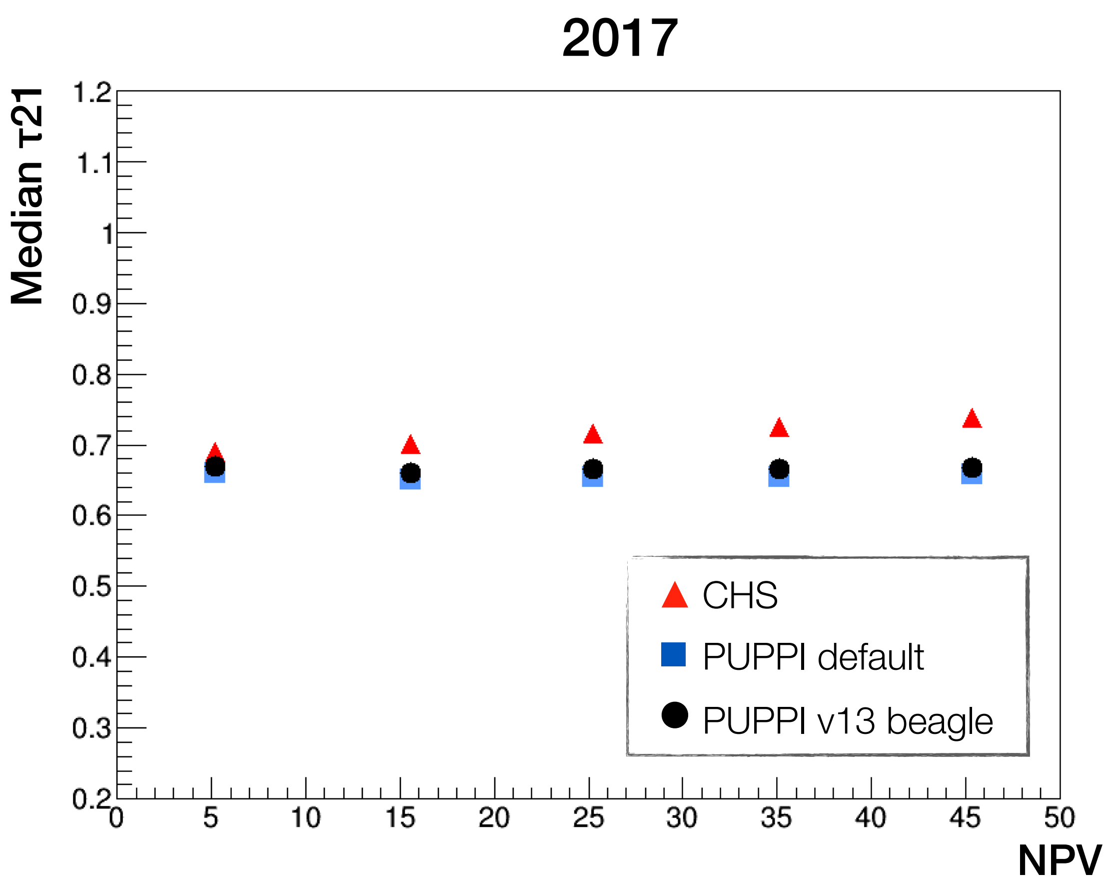


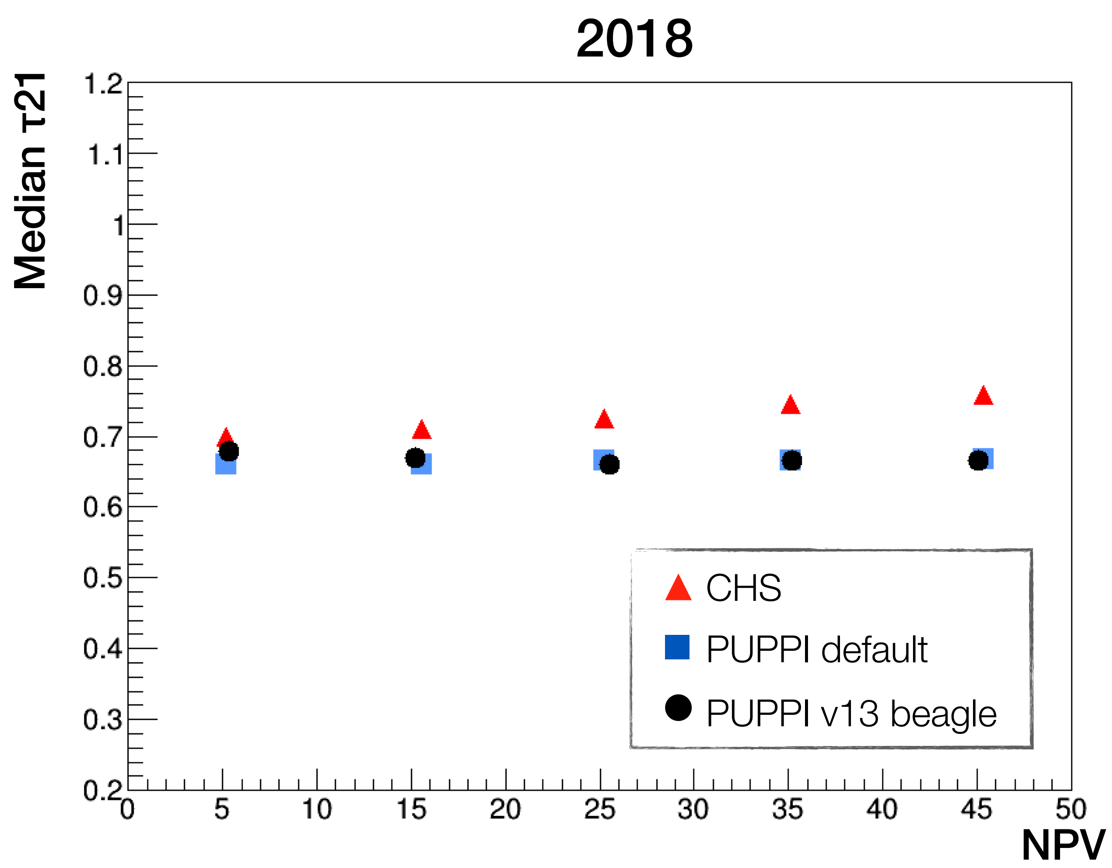
JER calculated with iterative gaussian fit

Median tau21

Median τ21







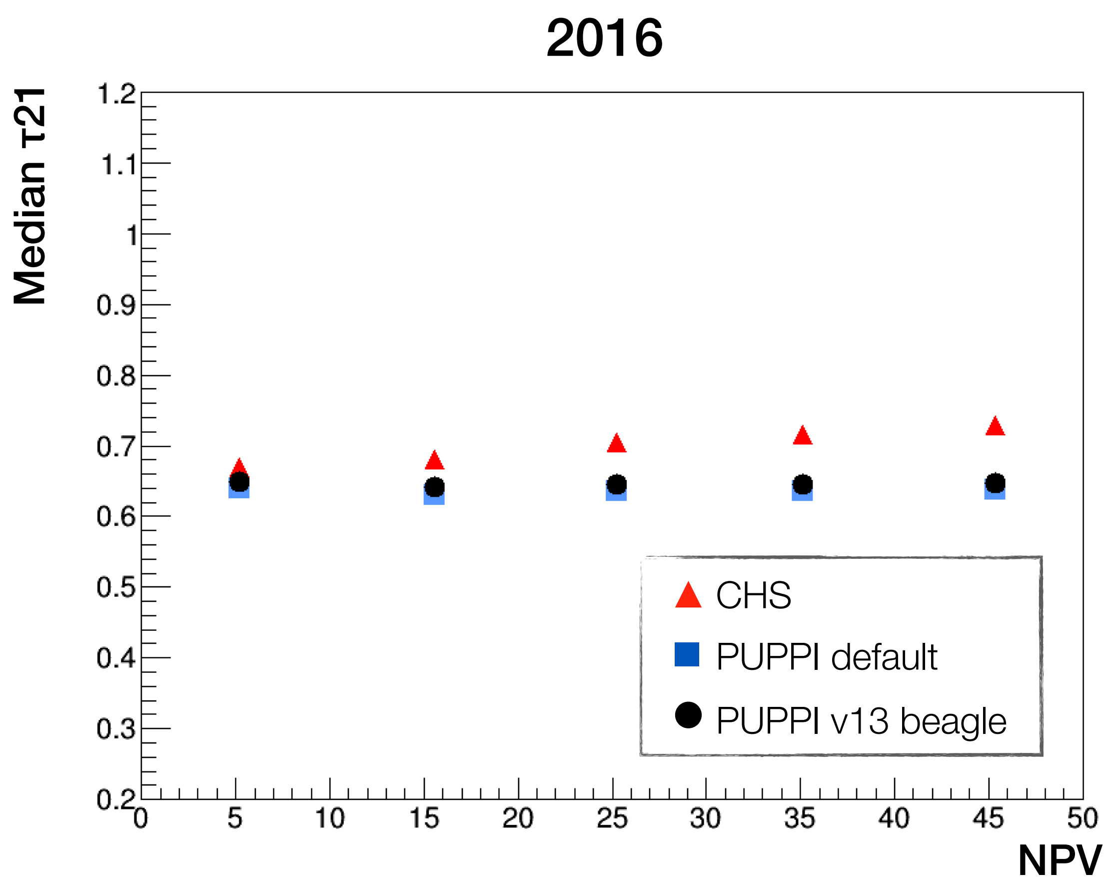
<u>**QCD sample**</u>

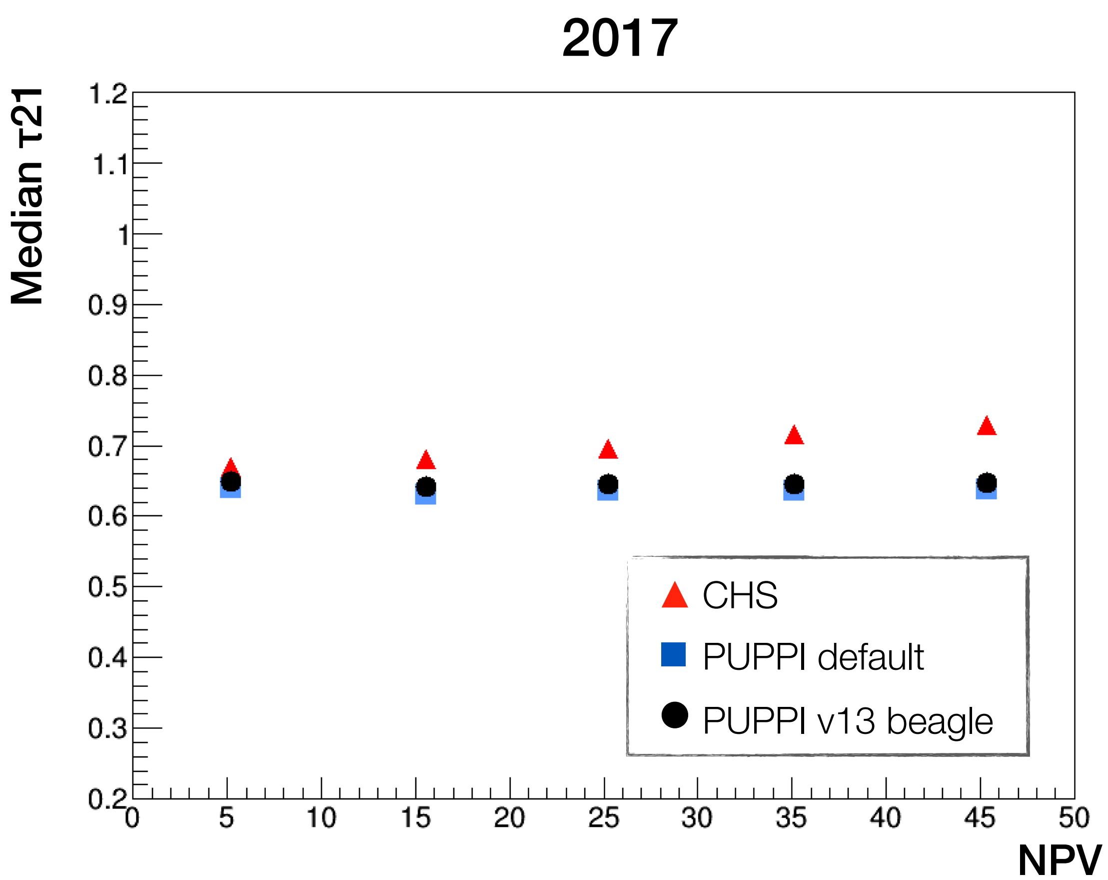
Median τ21 leading Ak8 jet

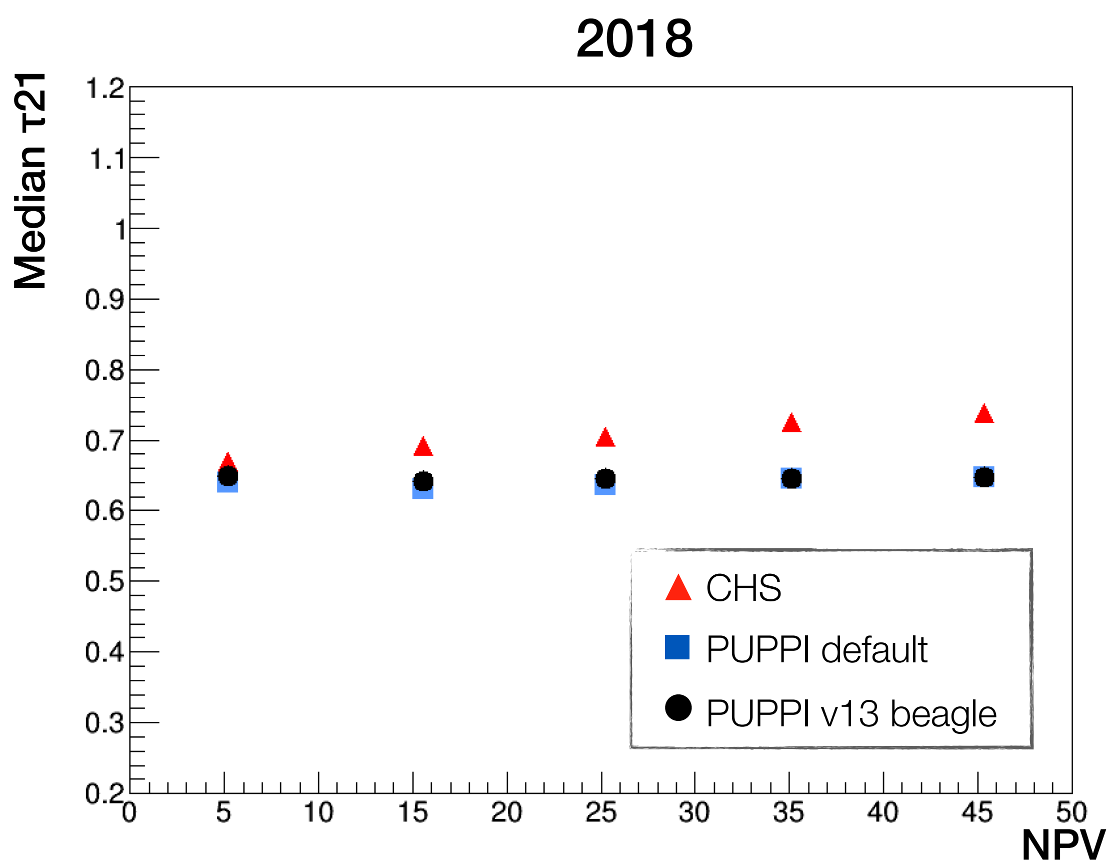
400 < p_T < 600 GeV

no SoftDrop Mass cut

Median τ21







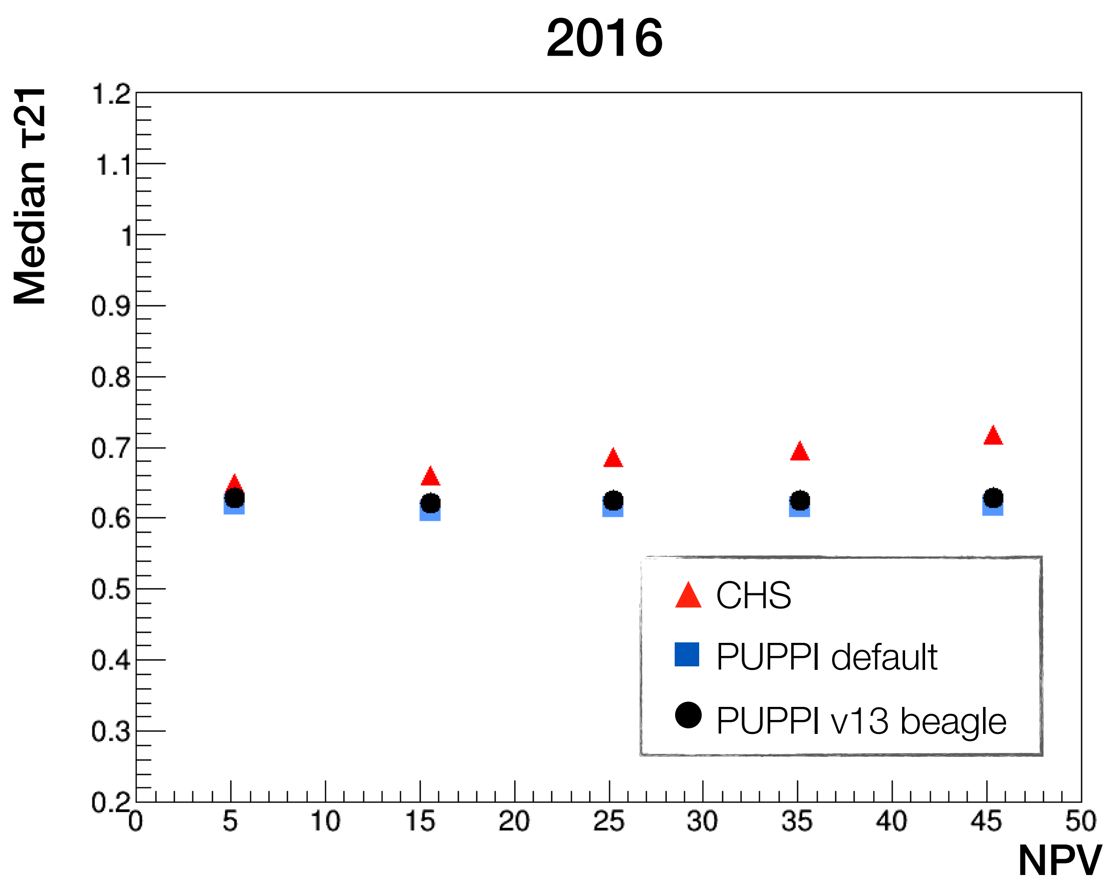
QCD sample

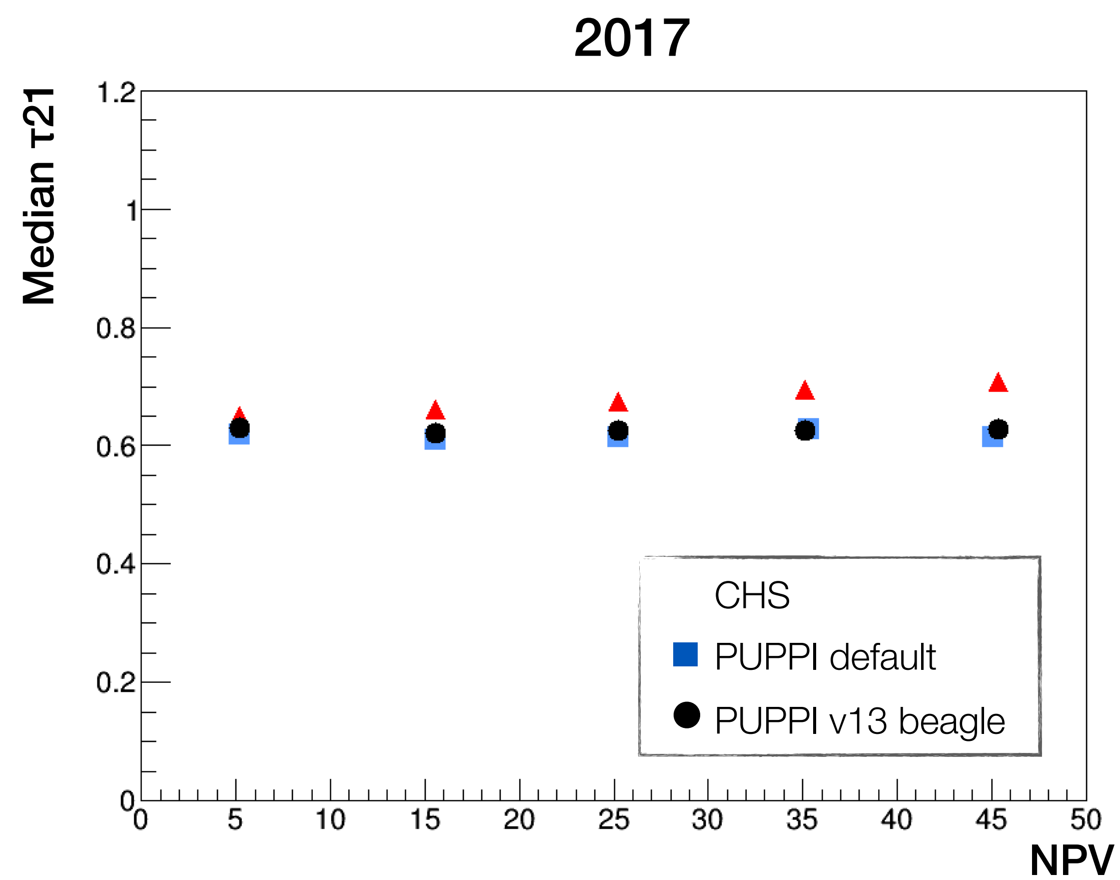
Median τ21 leading Ak8 jet

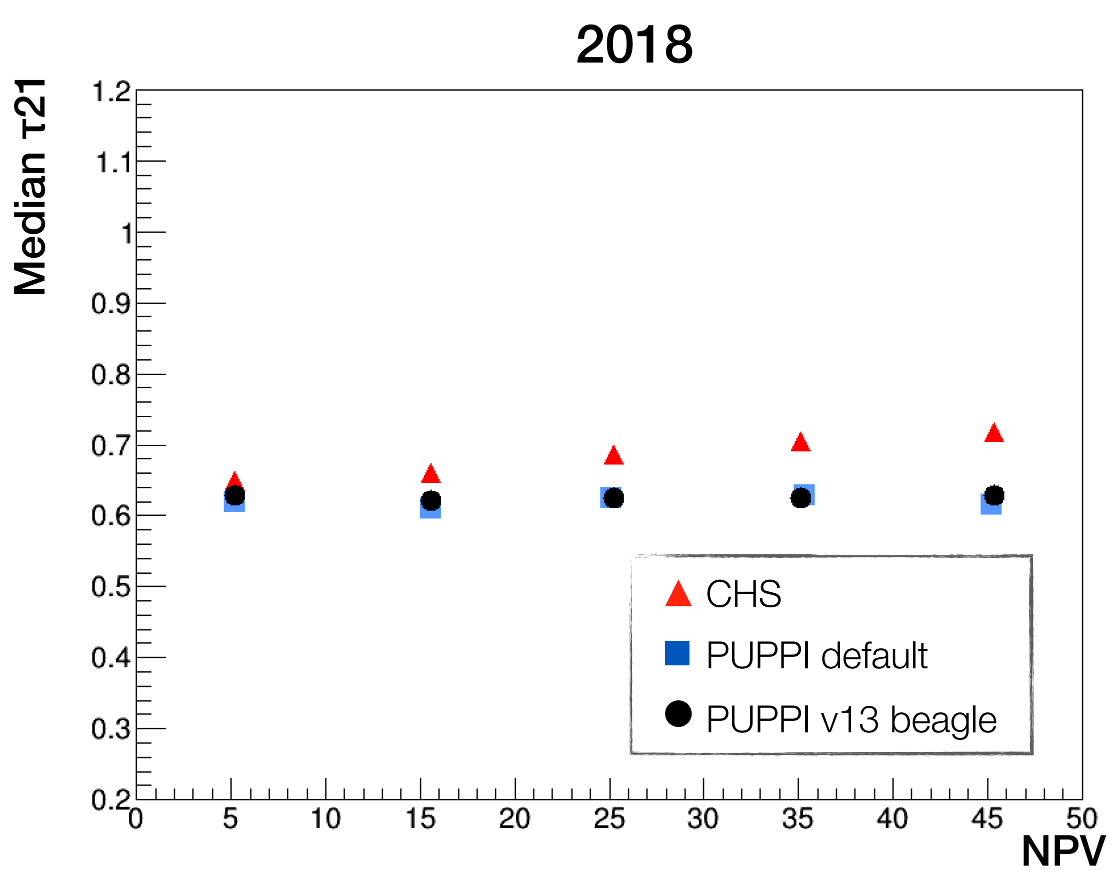
$600 < p_T < 800$ GeV

no SoftDrop Mass cut

Median τ21







<u>QCD sample</u>

Median τ21 leading Ak8 jet

800 < p_T < 1000 GeV

no SoftDrop Mass cut

Efficiency and Purity: calculation

Efficiency

- $p_T^{genjets} > 30$ GeV, $p_T^{recojets} > 20$ GeV
- Fill histogram with N_{PV} for each genjet
- Match each genjet to a recojet
- Fill histogram for each matched genjet with N_{PV}
- Divide the histograms

Purity

- $p_T^{recojets} > 30$ GeV, $p_T^{genjets} > 20$ GeV
- Fill histogram with N_{PV} for each recojet
- Match each recojet to a genjet
- Fill histogram for each matched recojet with N_{PV}
- Divide the histograms

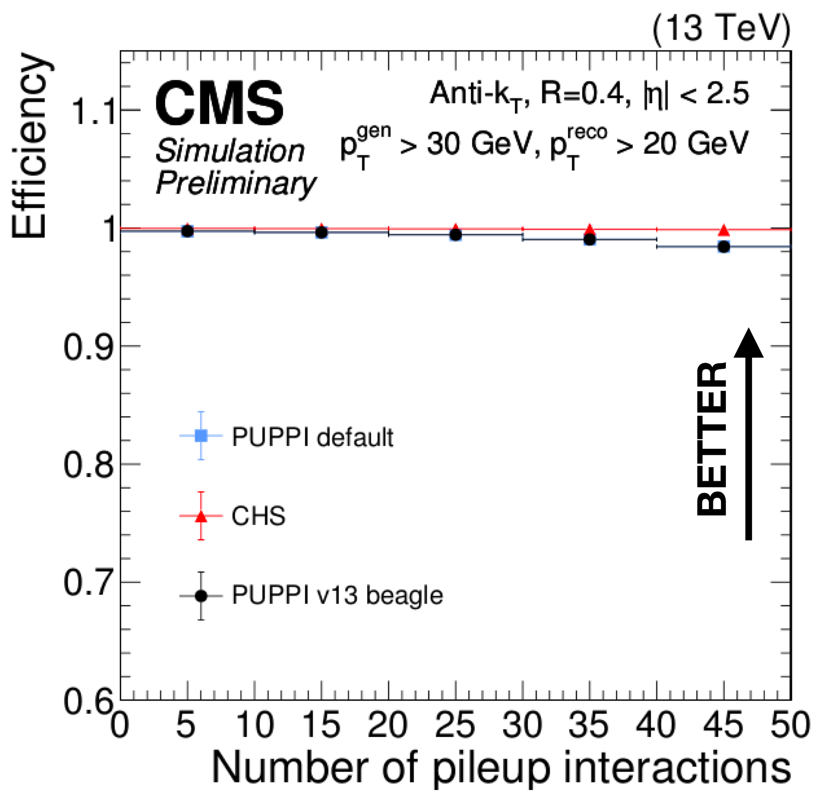
DY Dataset

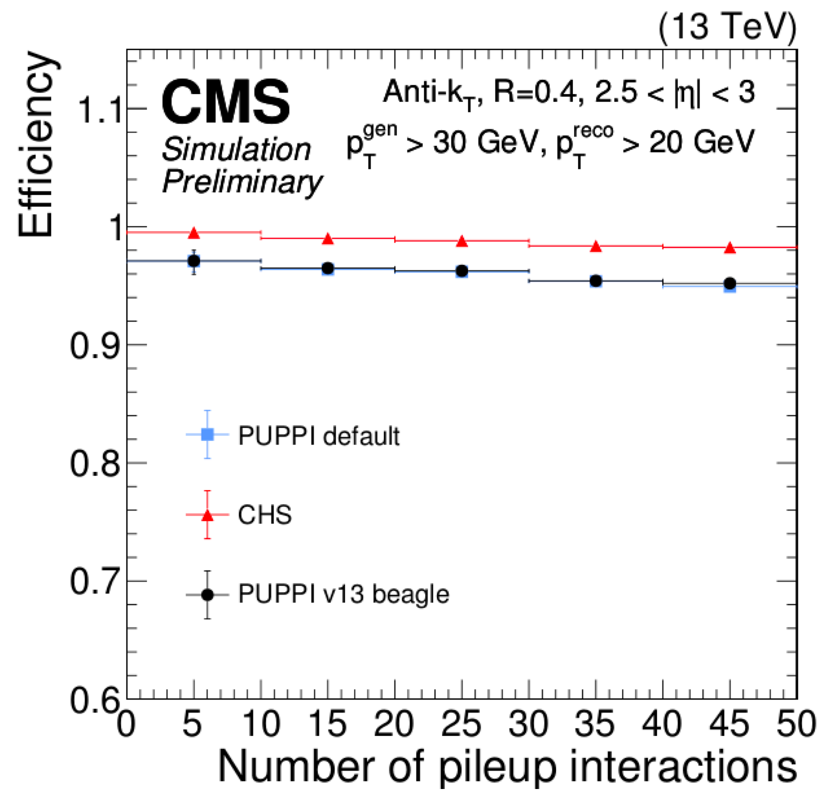
Year	DAS	CMSSW production release	Note
2016v3	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3_ext1-v2/MINIAODSIM	production: 80X reconstruction: 94X	Version v3, in QCD sample 2016 version v2
2017v2	/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIIFall17MiniAODv2-PU2017_12Apr2018_new_pmx_94X_mc2017_realistic_v14-v1/MINIAODSIM	production: 92X reconstruction: 94X	
2018	/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM	production: 102X reconstruction: 102X	Coming soon

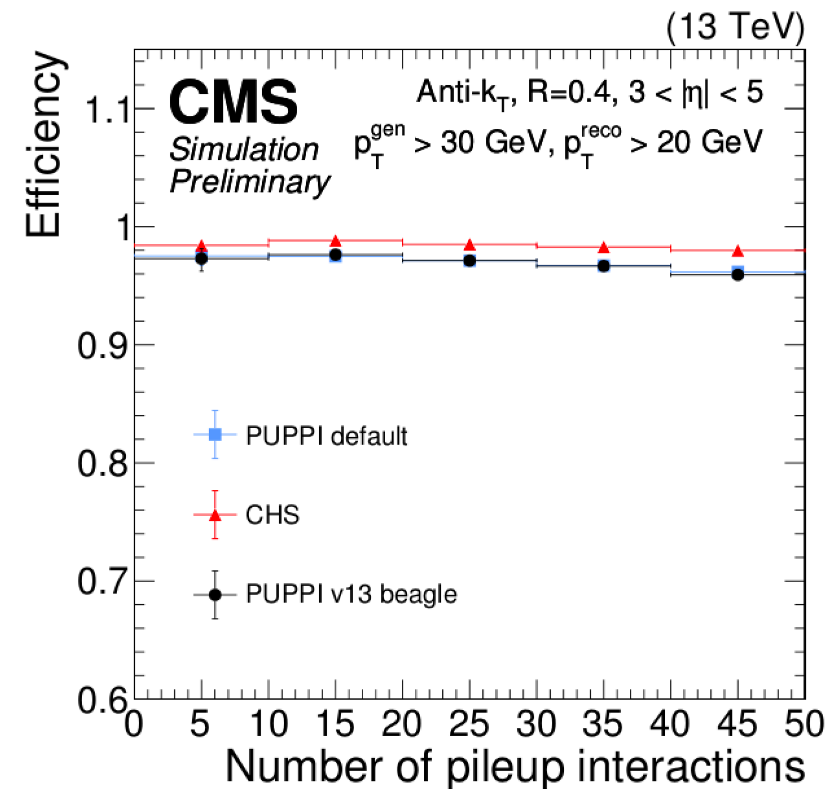
PR tune v13 beagle

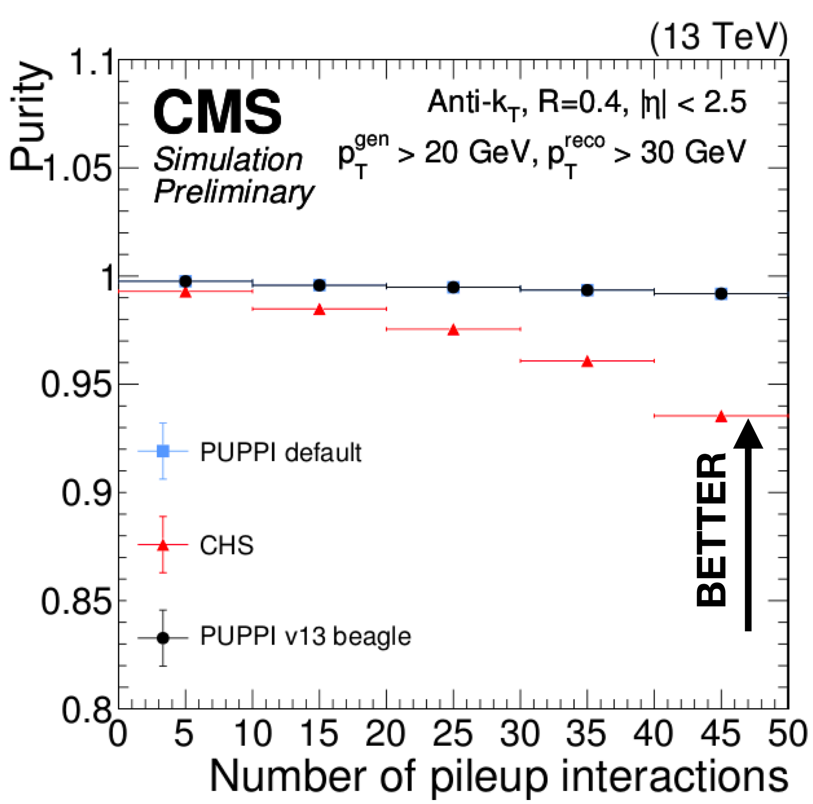
Efficiency & Purity - 2016

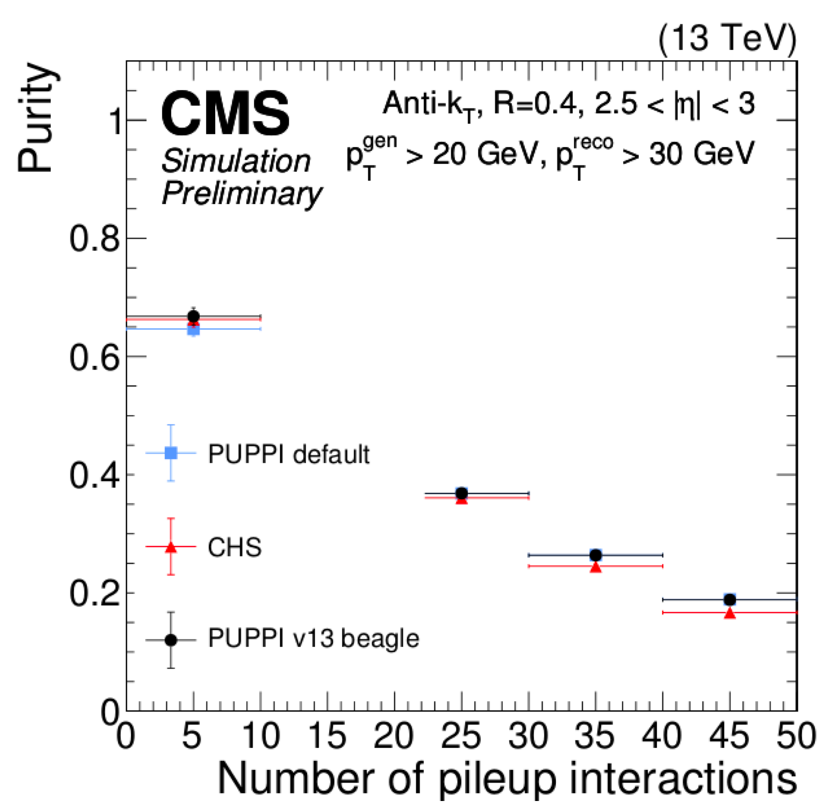
<u>DY sample</u>

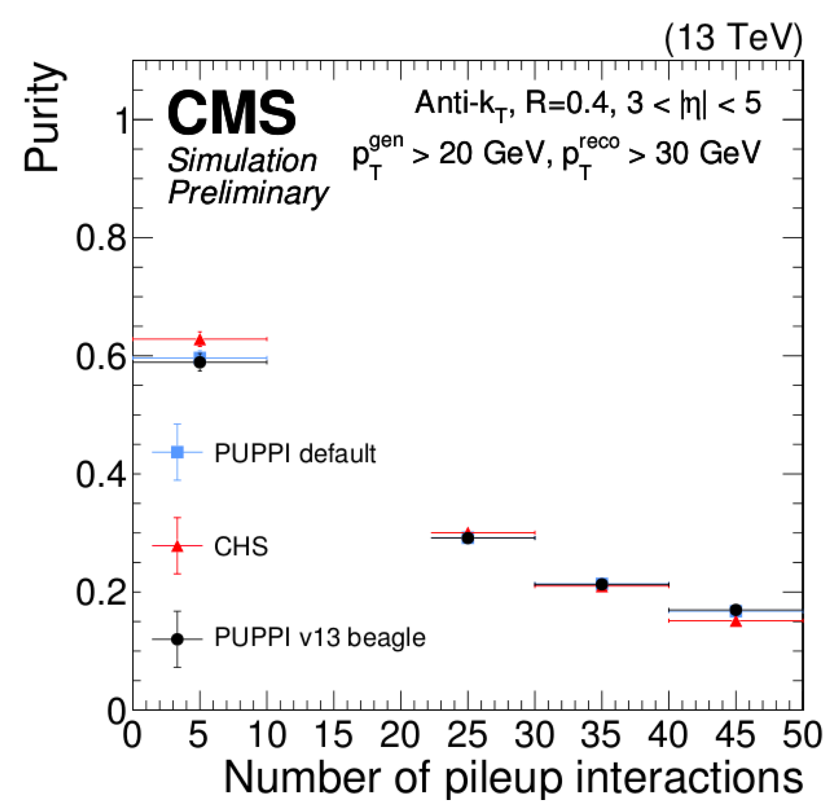












Efficiency & Purity - 2017

DY sample