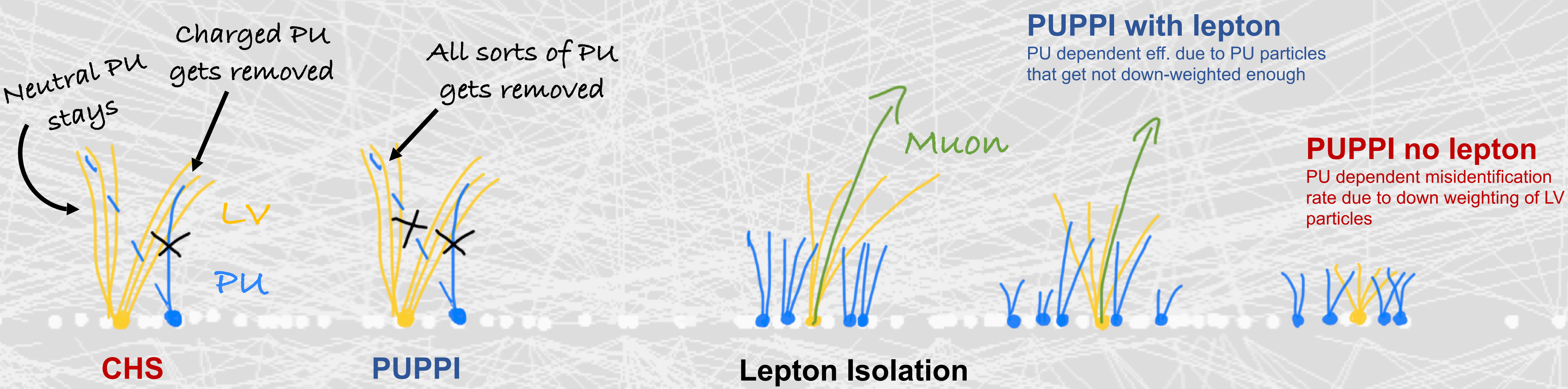
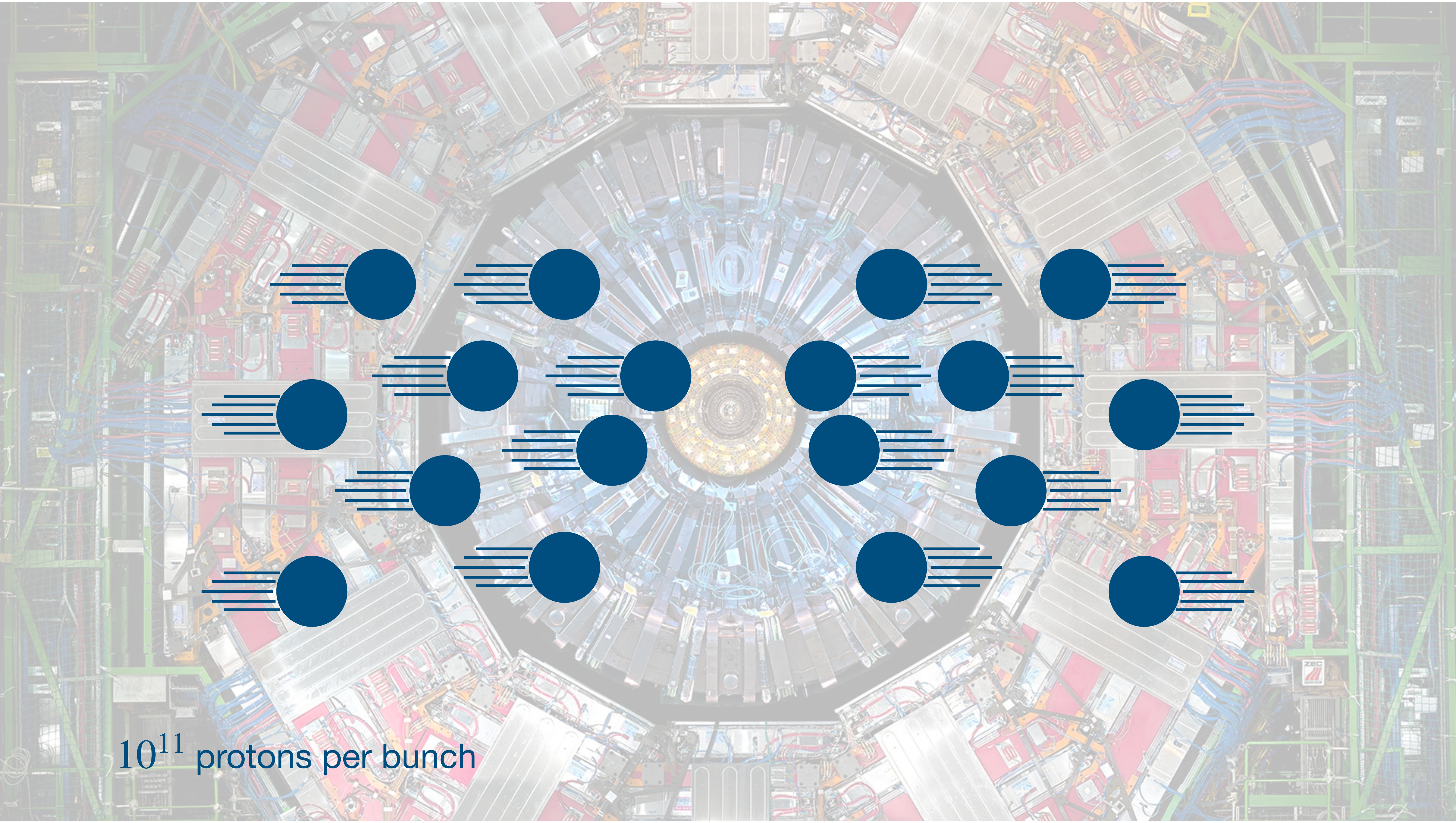
How CMS weeds out particles that pile up

Anna Benecke
UCLouvain

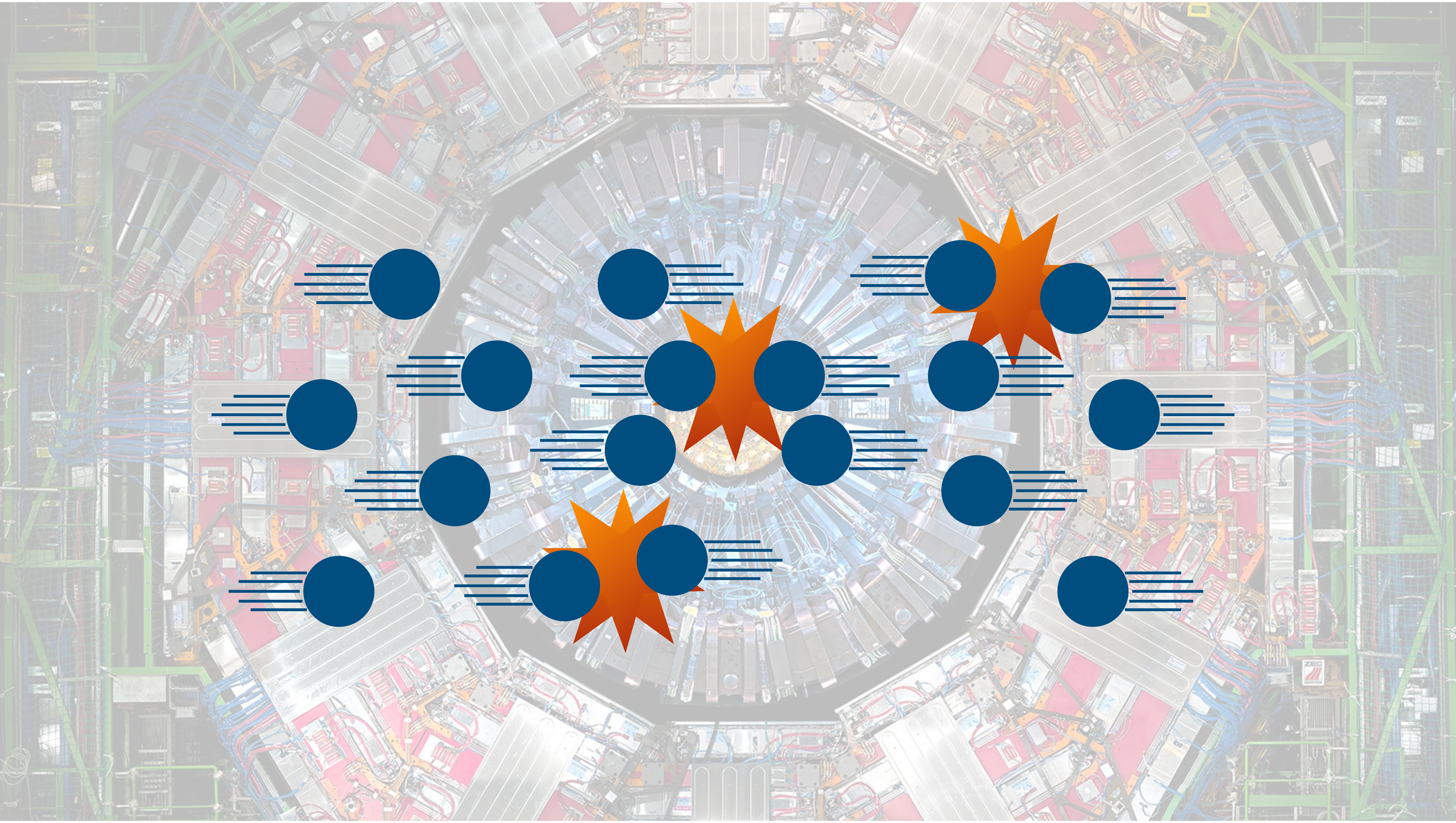


Proton-Proton collisions @ LHC

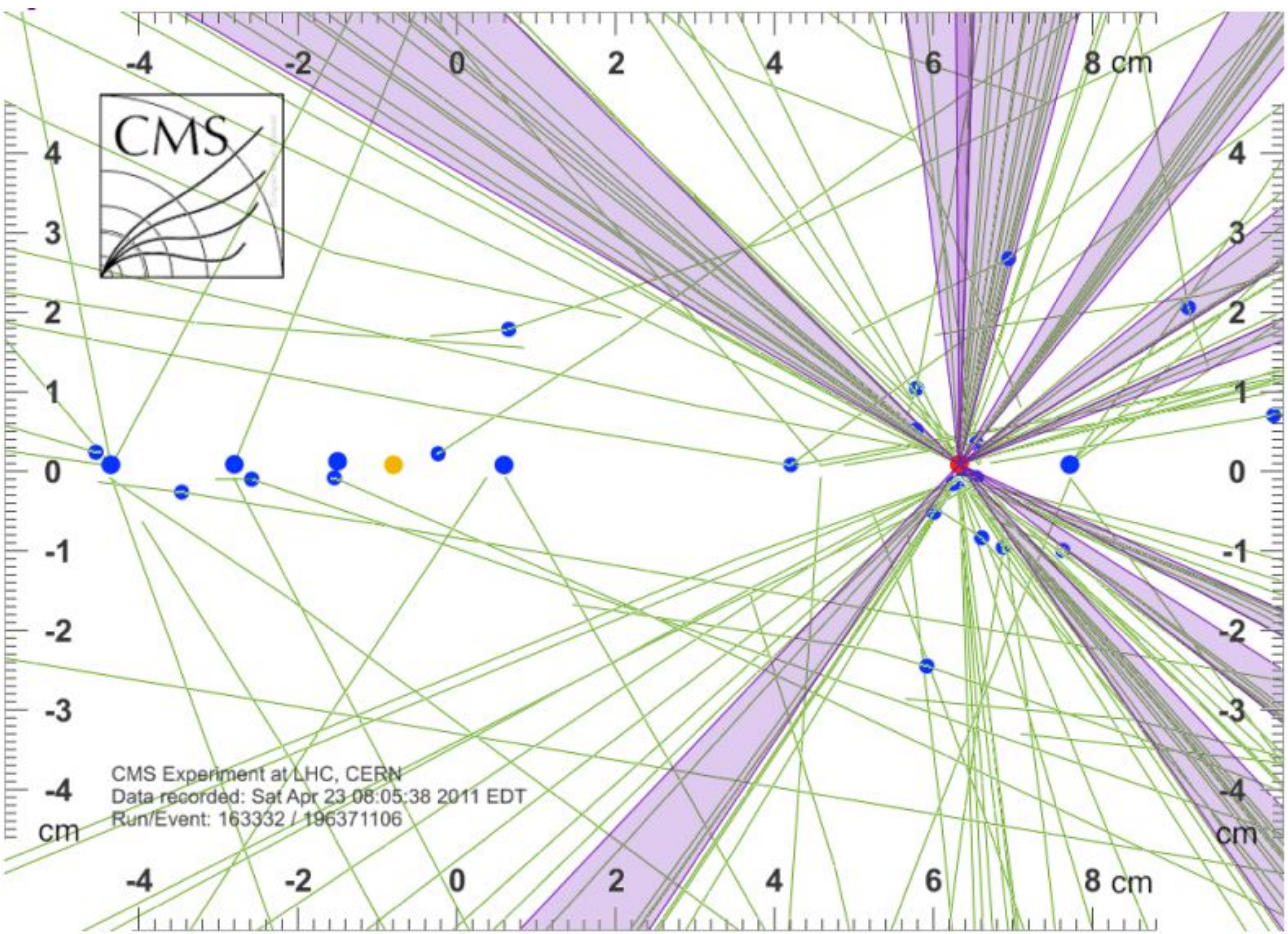


10^{11} protons per bunch

Proton-Proton collisions @ LHC



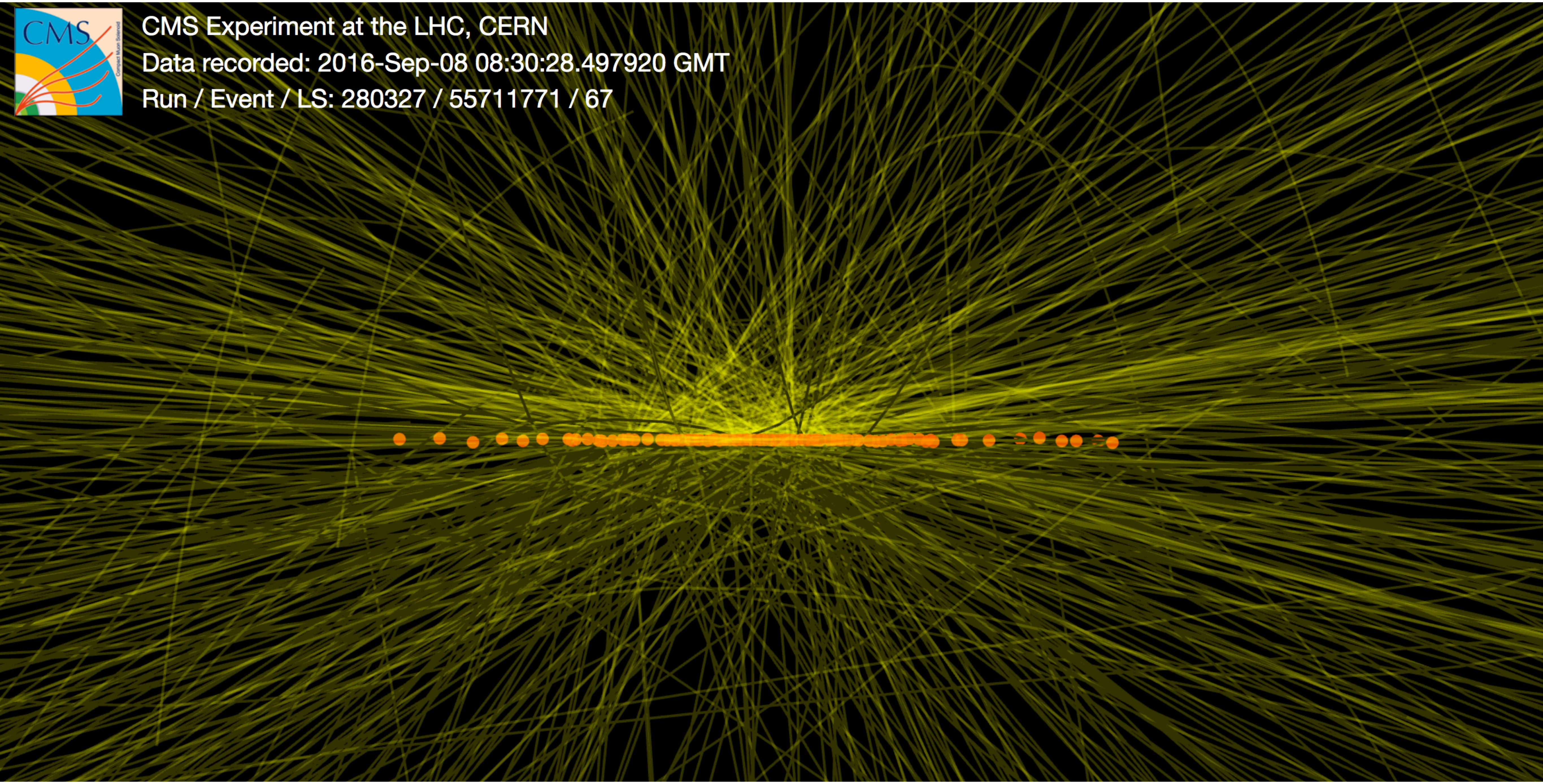
Proton-Proton collisions @ LHC



http://cds.cern.ch/record/1357882

Pileup adds additional energy to the whole detector

Proton-Proton collisions @ LHC



Pileup adds additional energy to the whole detector

More data, more pileup

$$N_{\text{Events}} = \sigma \int L dt$$

frequency
40 MHz

Number of protons

$$L = \frac{f N_1 N_2}{4\pi \sigma_x \sigma_y}$$

Size of the beam

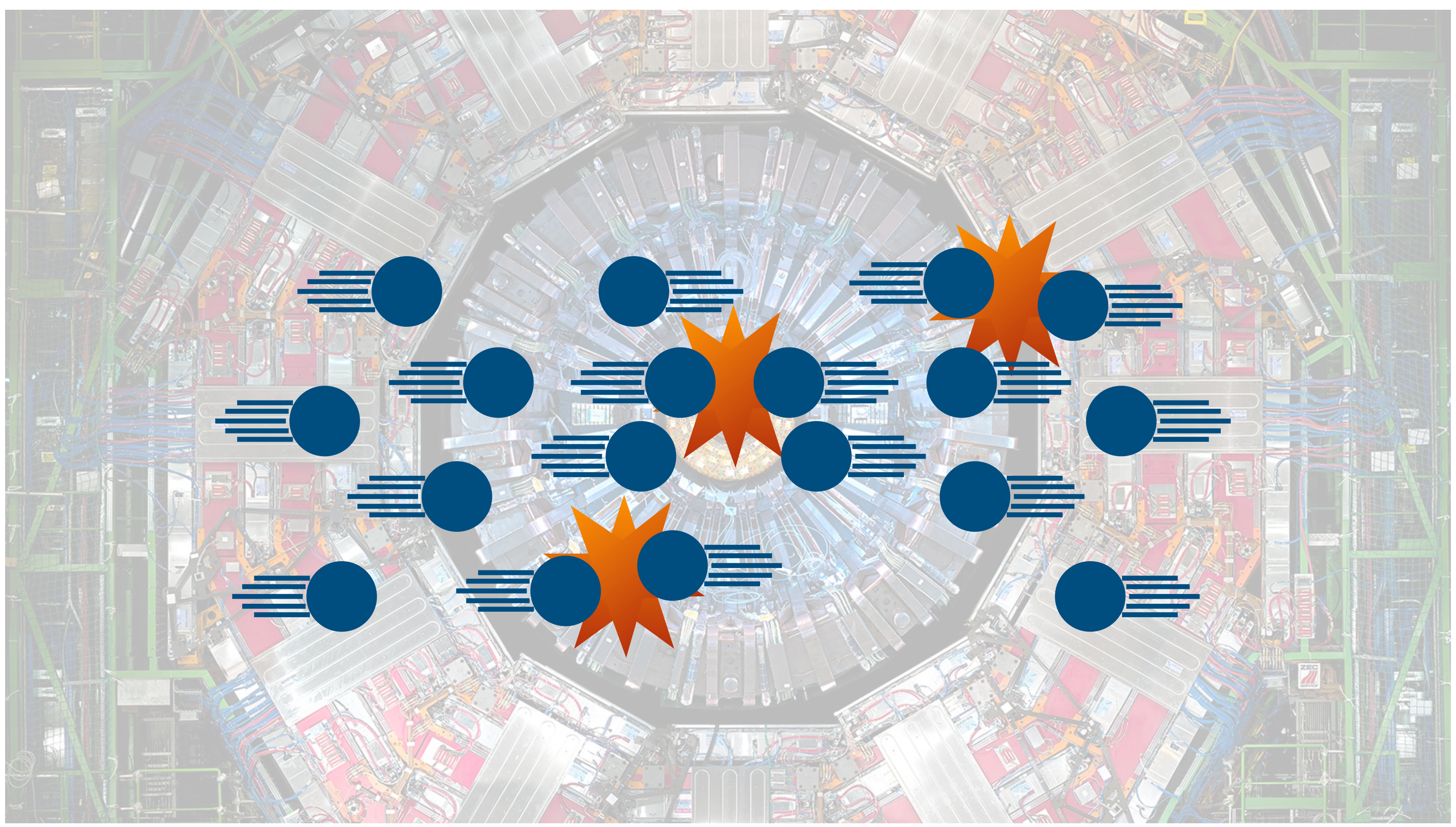
More data, more pileup

$$N_{\text{Events}} = \sigma \int L dt \qquad L = \frac{f N_1 N_2}{4\pi\sigma_x\sigma_y}$$

frequency
40 MHz

Number of protons

Size of the beam



More data, more pileup

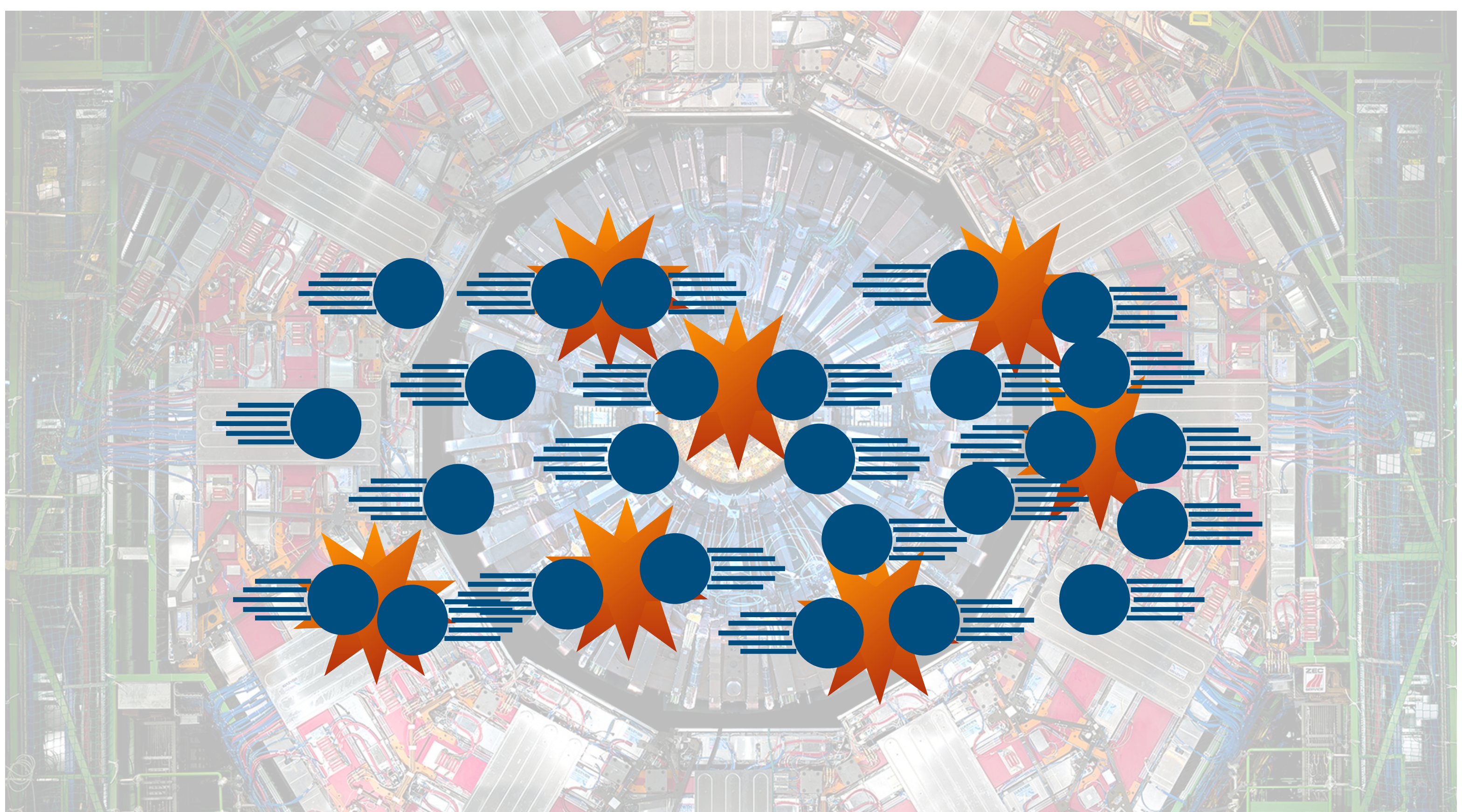
$$N_{\text{Events}} = \sigma \int L dt$$

frequency
40 MHz

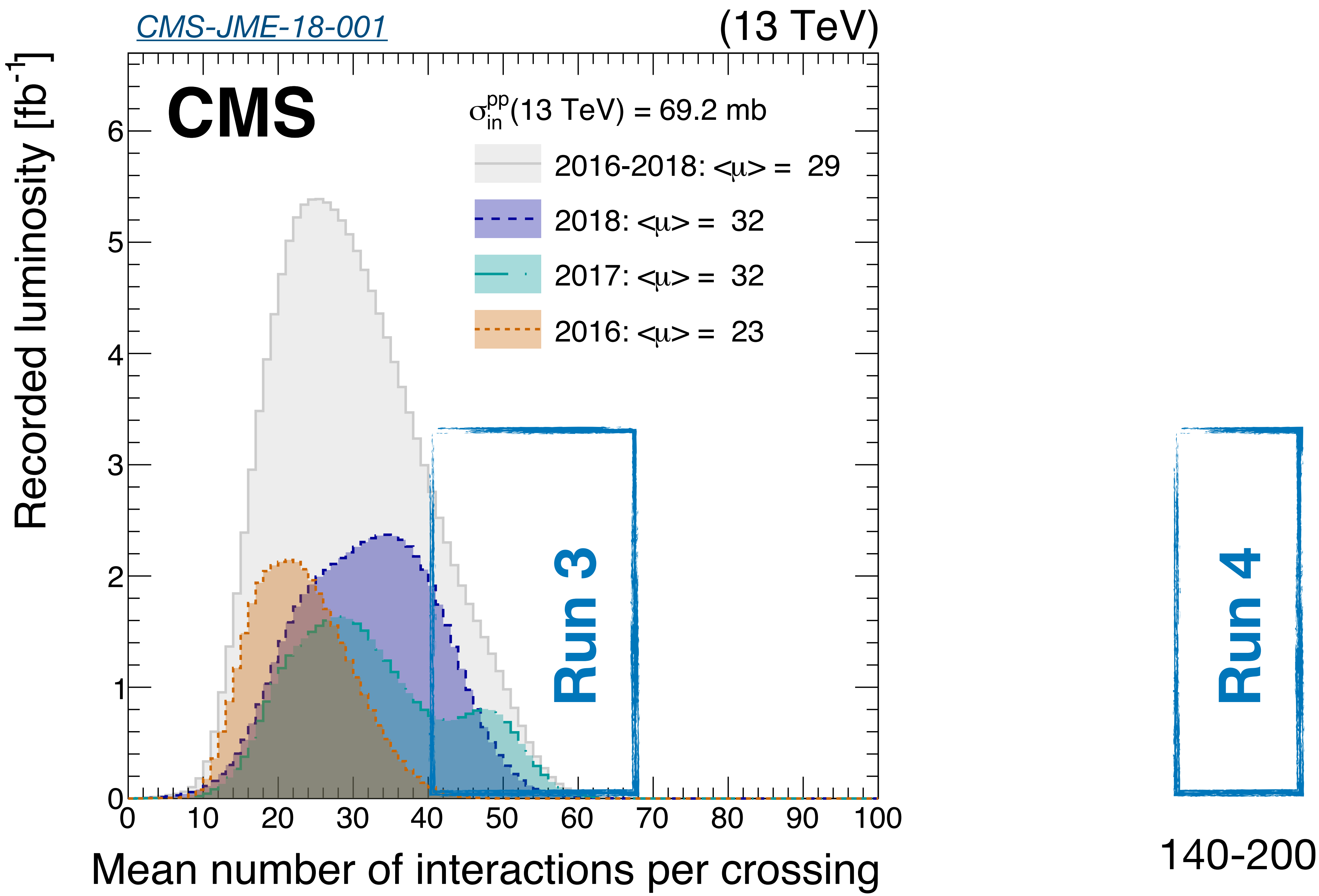
Number of protons

$$L = \frac{f N_1 N_2}{4\pi \sigma_x \sigma_y}$$

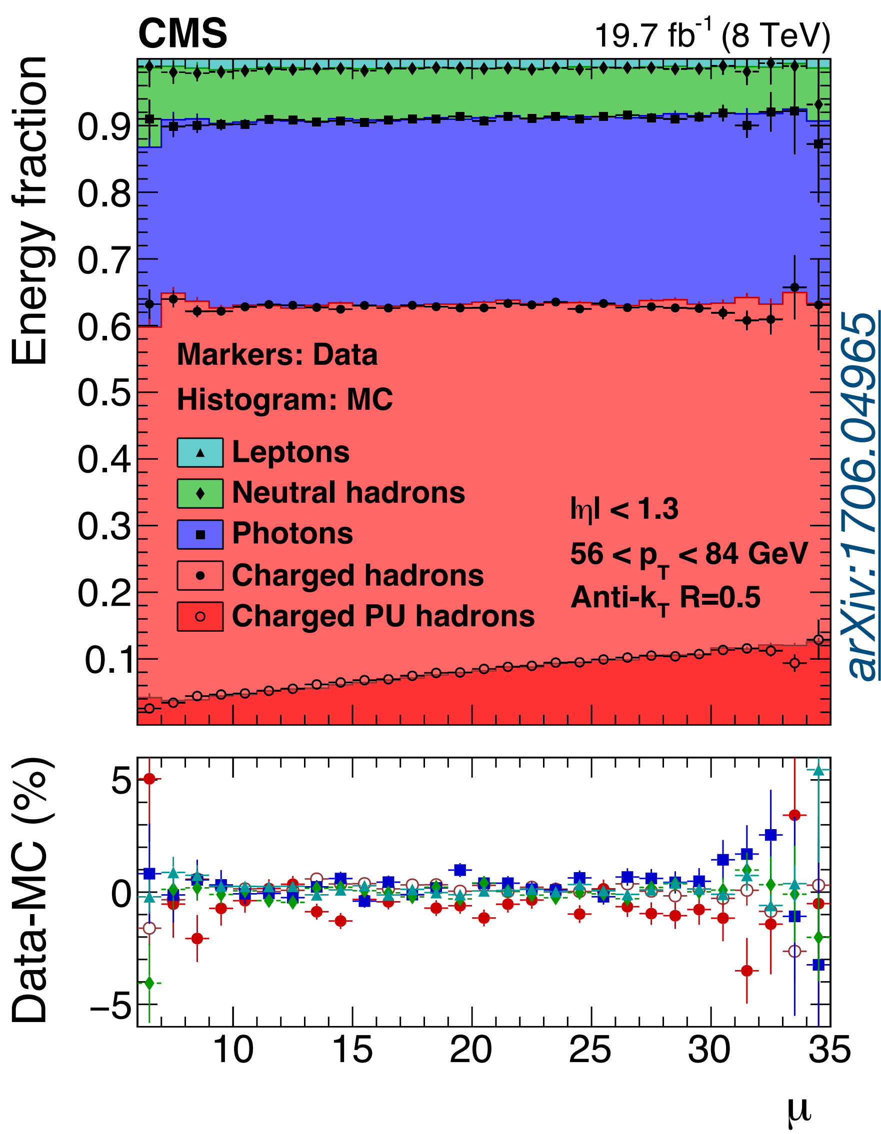
Size of the beam

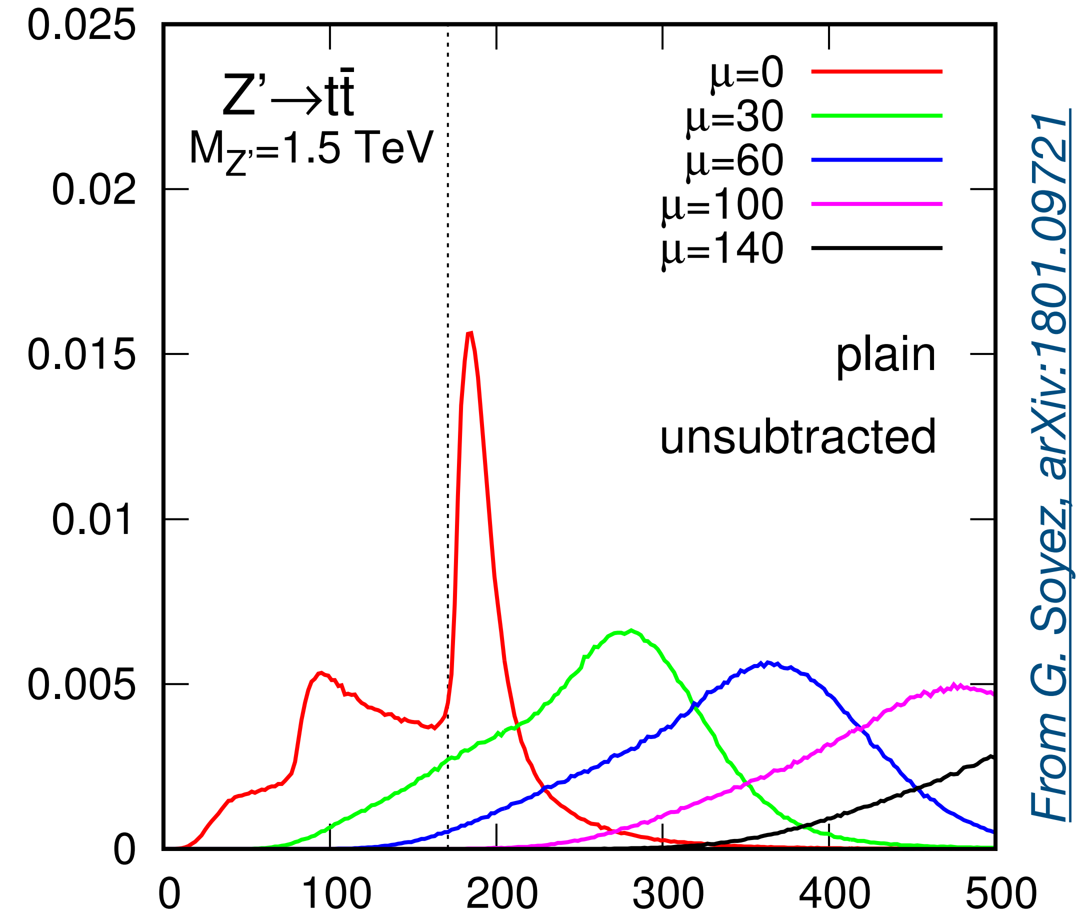


More data, more pileup



Challenges with pileup

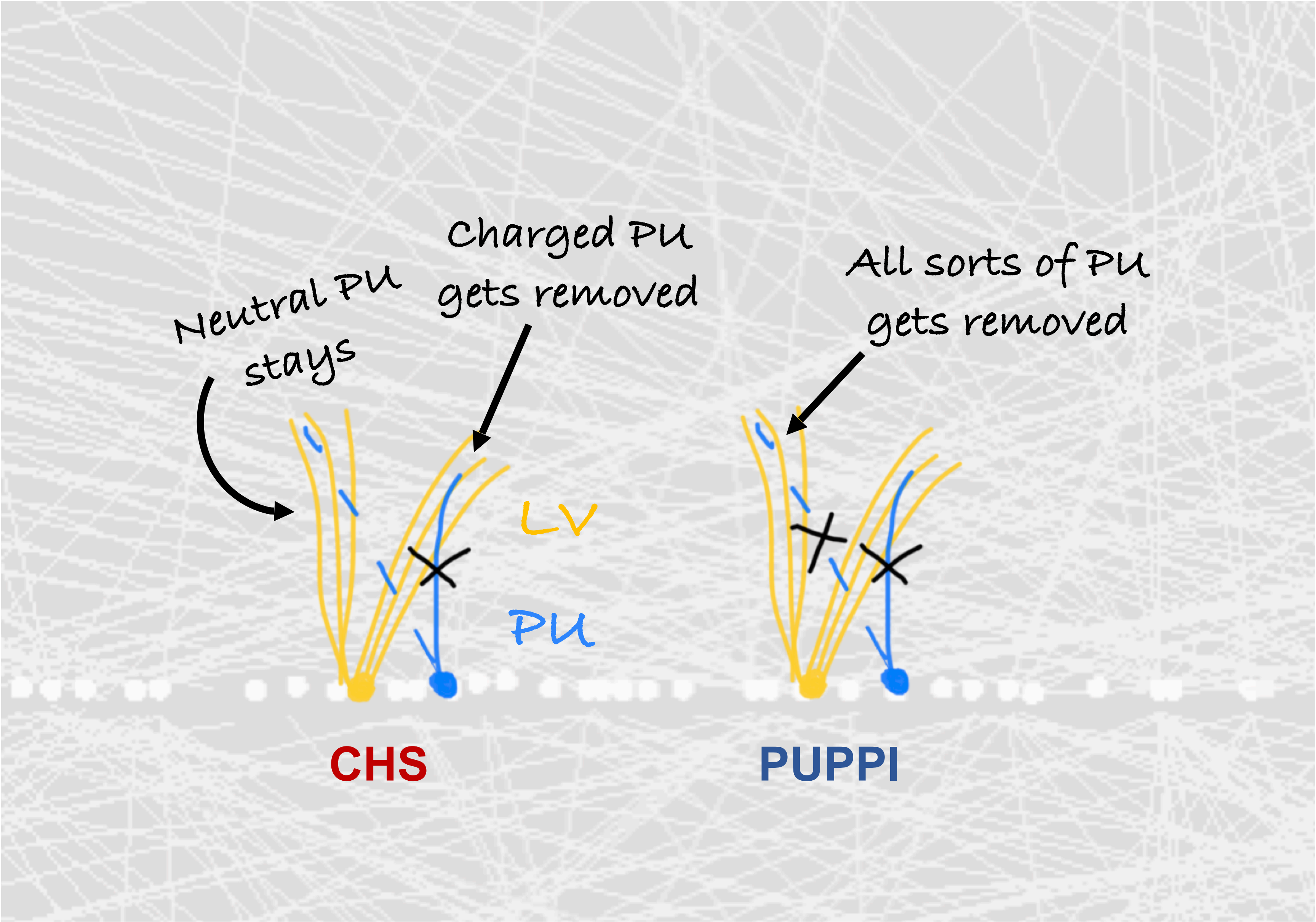




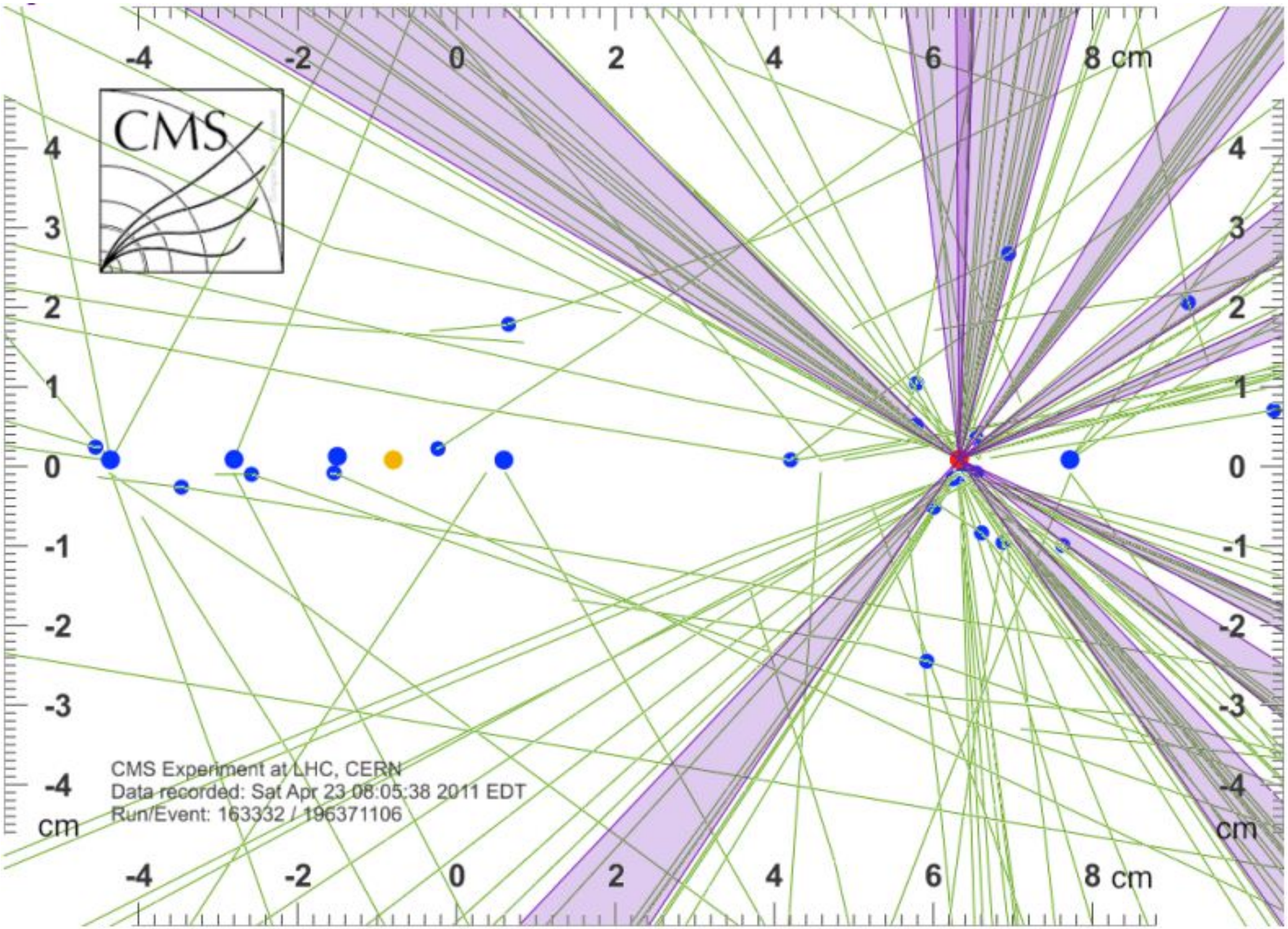
(a) raw, ungroomed jets

PU affects jet substructure, jet counting, lepton isolation…

Pileup mitigation in CMS

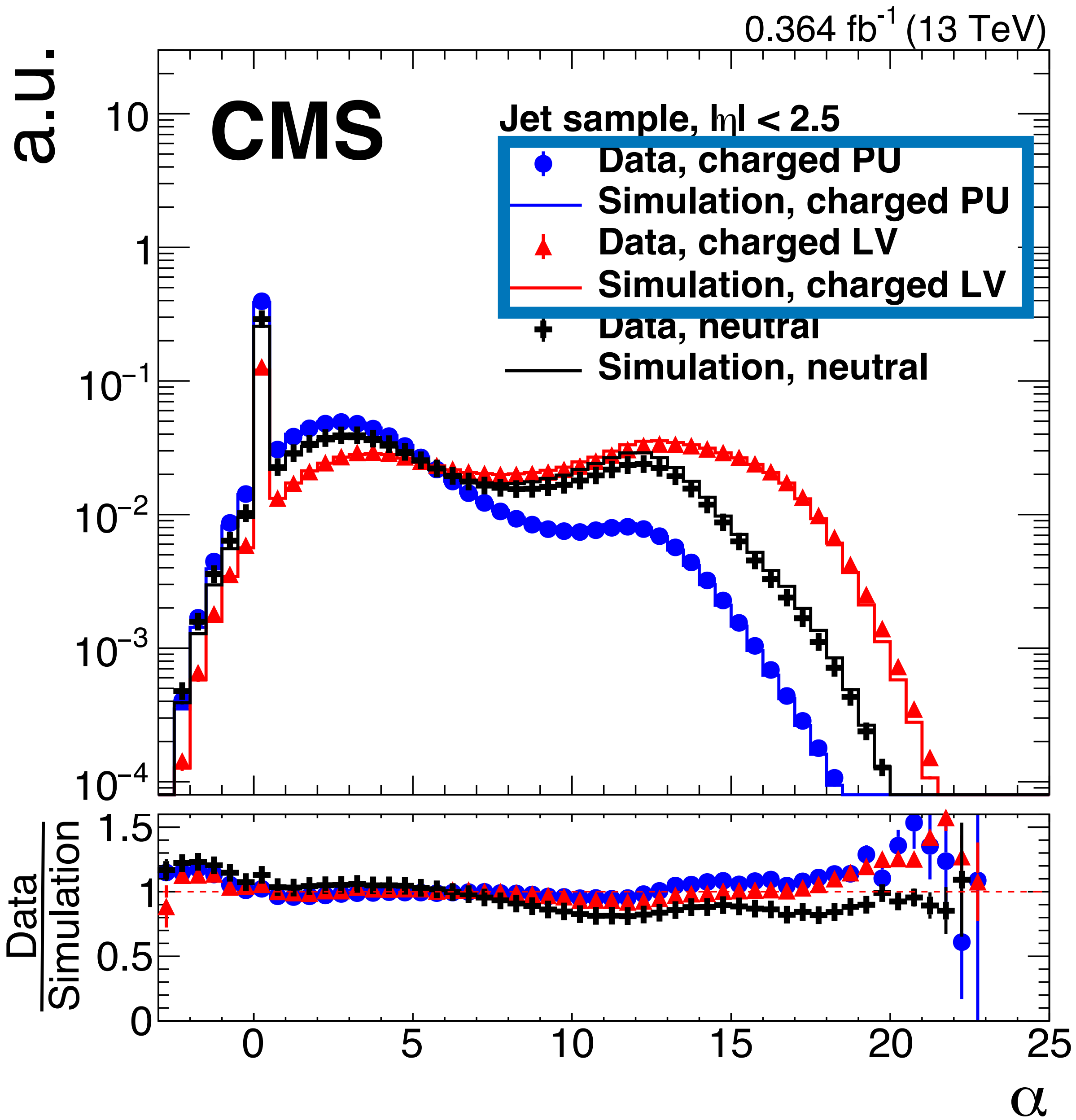


PUPPI in Detail

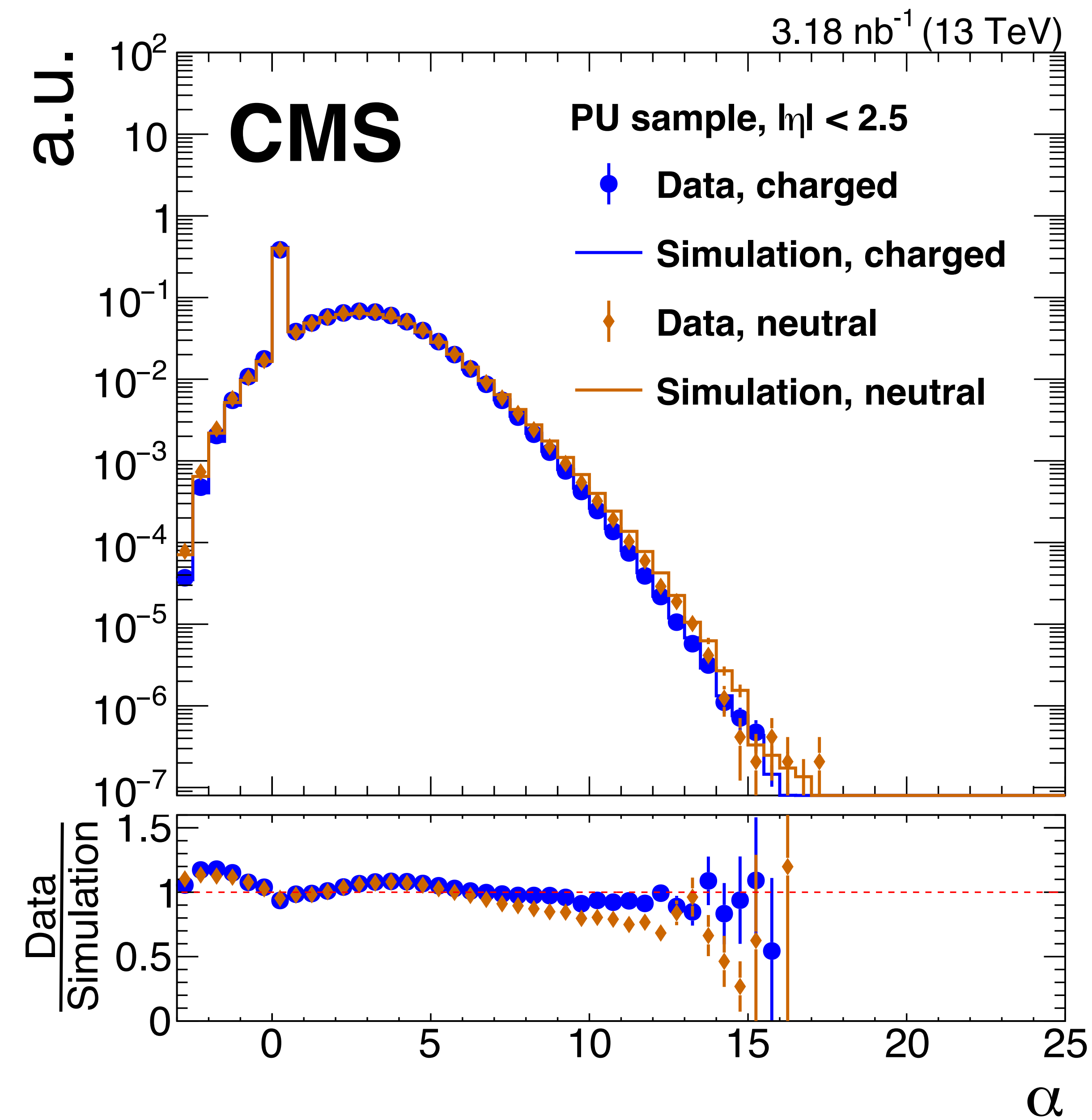


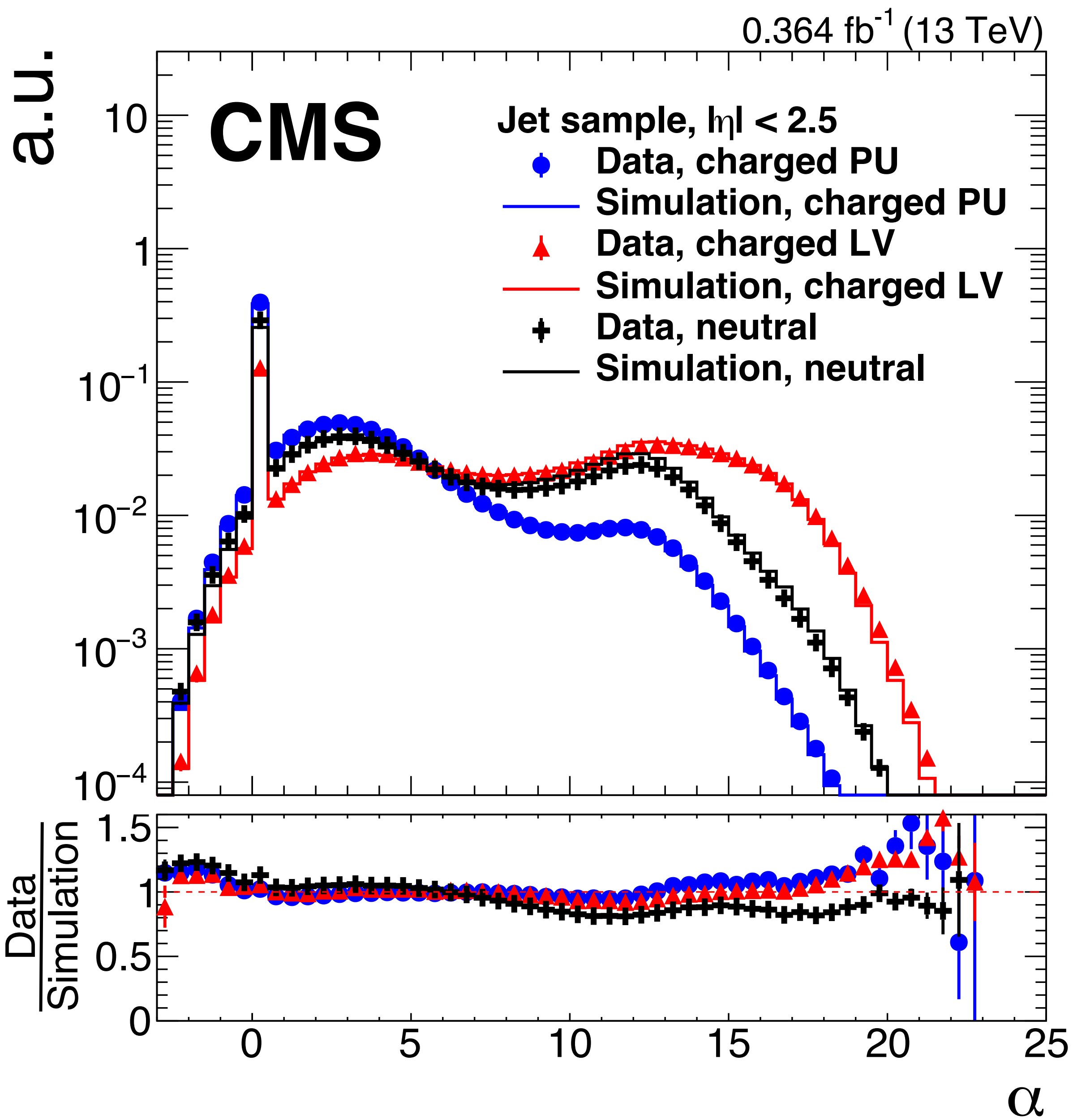
http://cds.cern.ch/record/1357882

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



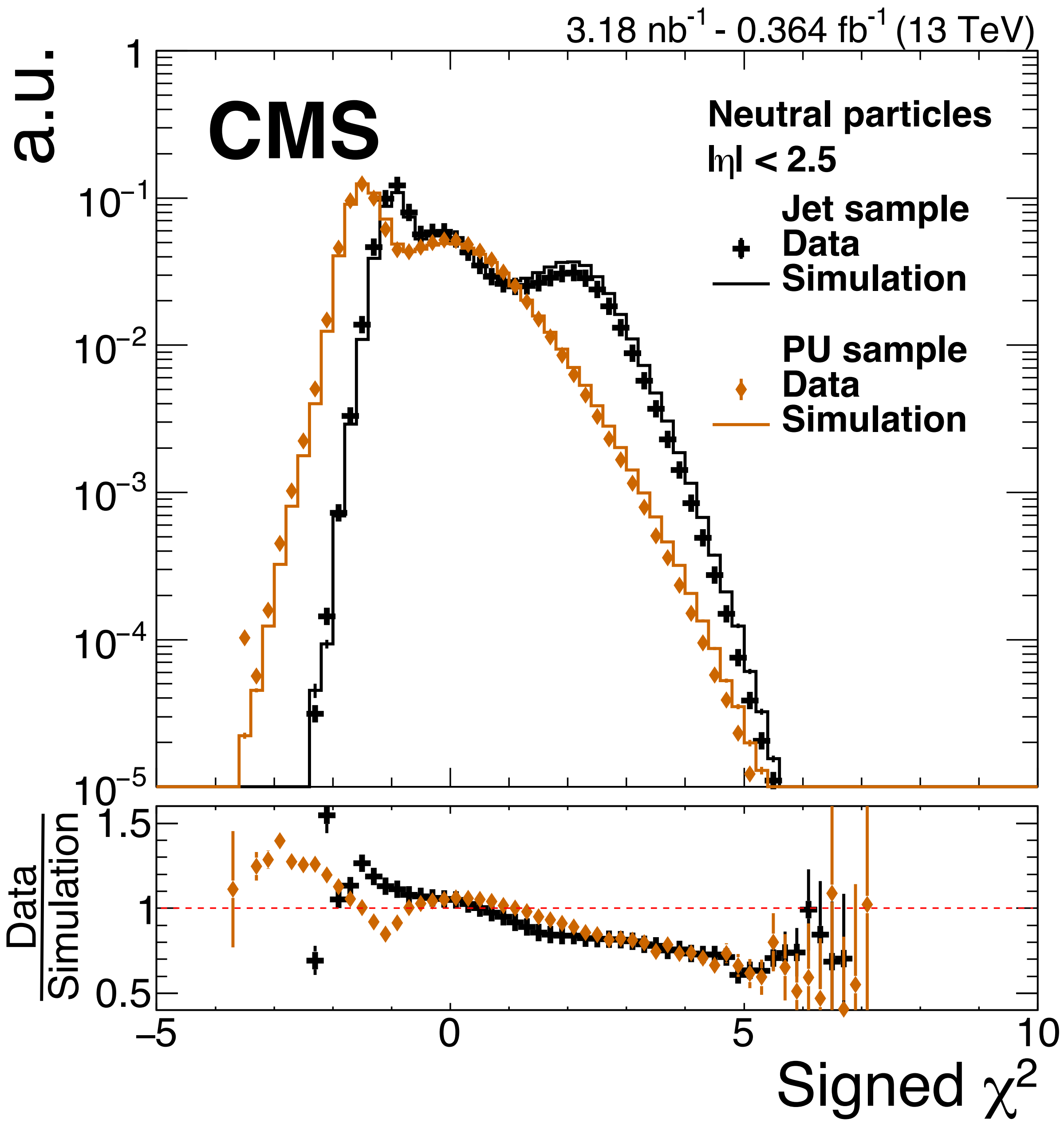
$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{\mathrm{T}j}}{\Delta R_{ij}} \right)^2 \begin{cases} \text{for } |\eta_i| < 2.5, & j \text{ are charged particles from leading vertex} \\ \text{for } |\eta_i| > 2.5, & j \text{ are all kinds of reconstructed particles} \end{cases}$$





1. Calculate Median and RMS of charged PU shape (blue)

$$\bar{\alpha}_{PU},\ RMS_{PU}$$

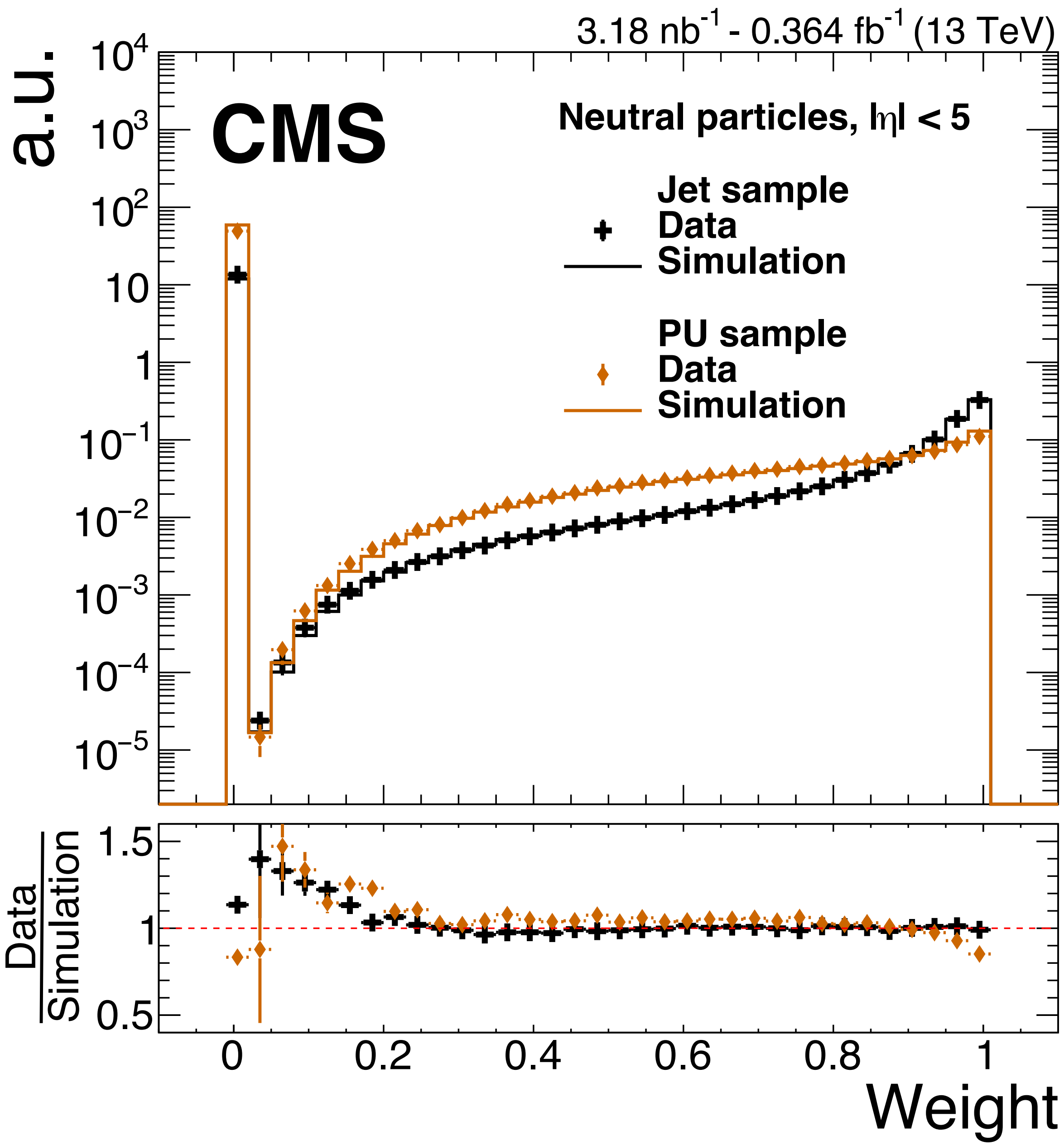


1. Calculate Median and RMS of charged PU shape

$$\bar{\alpha}_{PU},\ RMS_{PU}$$

2. For each particle calculate

$$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})\,|\,\alpha_i - \bar{\alpha}_{PU}\,|}{RMS_{PU}^2}$$



1. Calculate Median and RMS of charged PU shape

$$\bar{\alpha}_{PU},\ RMS_{PU}$$

2. For each particle calculate

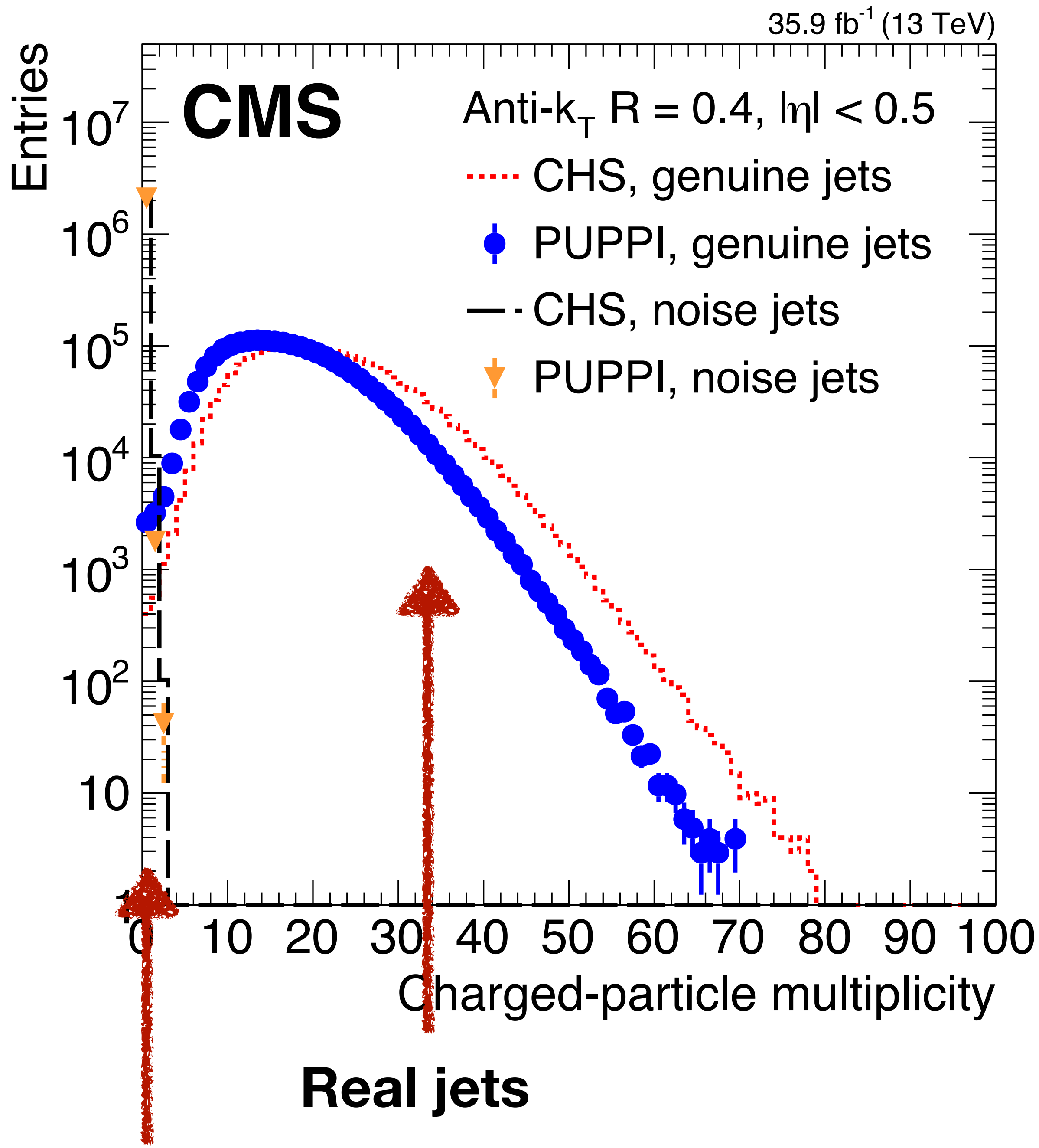
$$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})\,|\,\alpha_i - \bar{\alpha}_{PU}\,|}{RMS_{PU}^2}$$

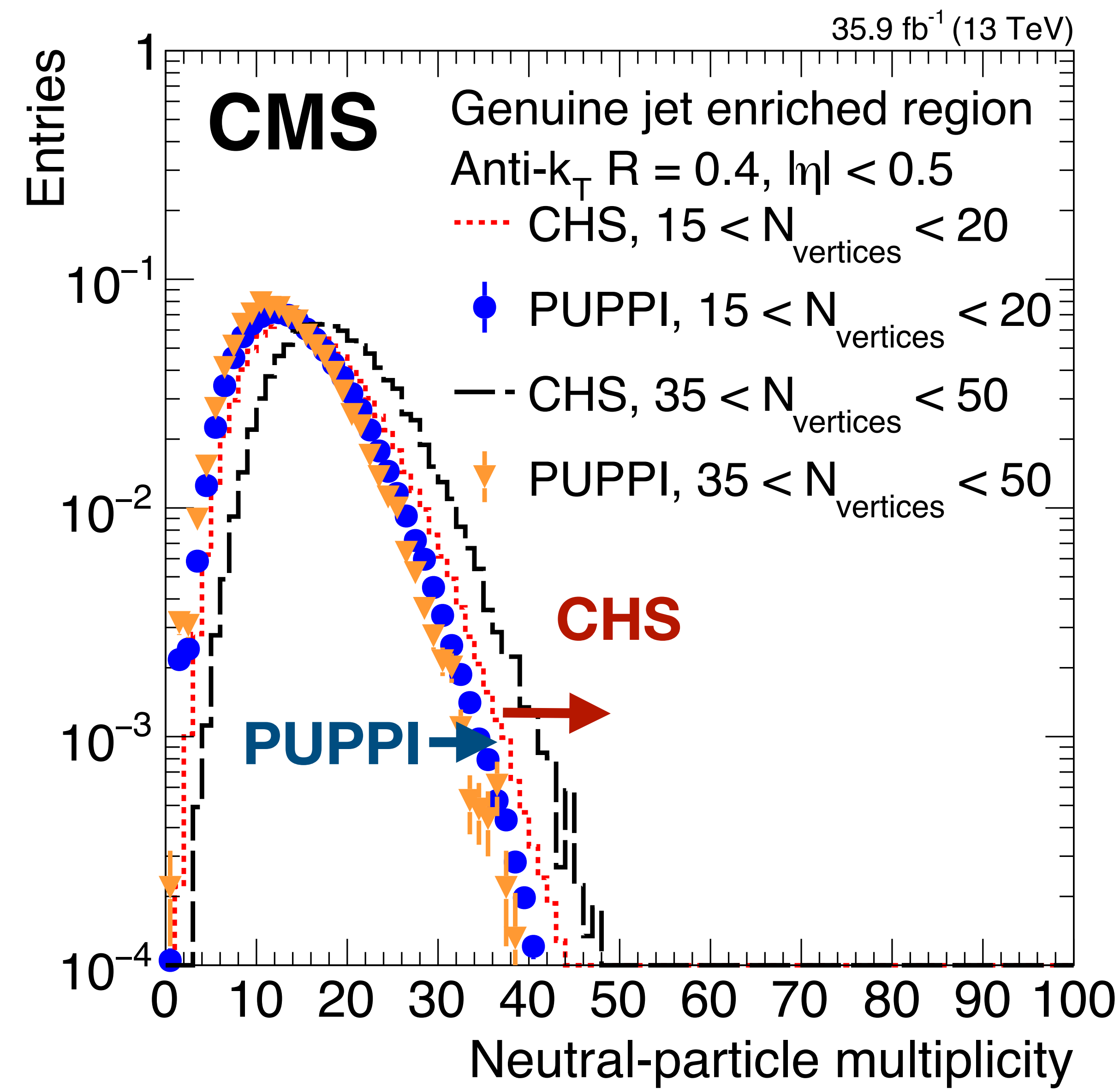
3. Assign a weight to each particle

$$w_i = F_{\chi^2, NDF=1}(\chi_i^2)$$

Pileup ⟶ Hard scattering

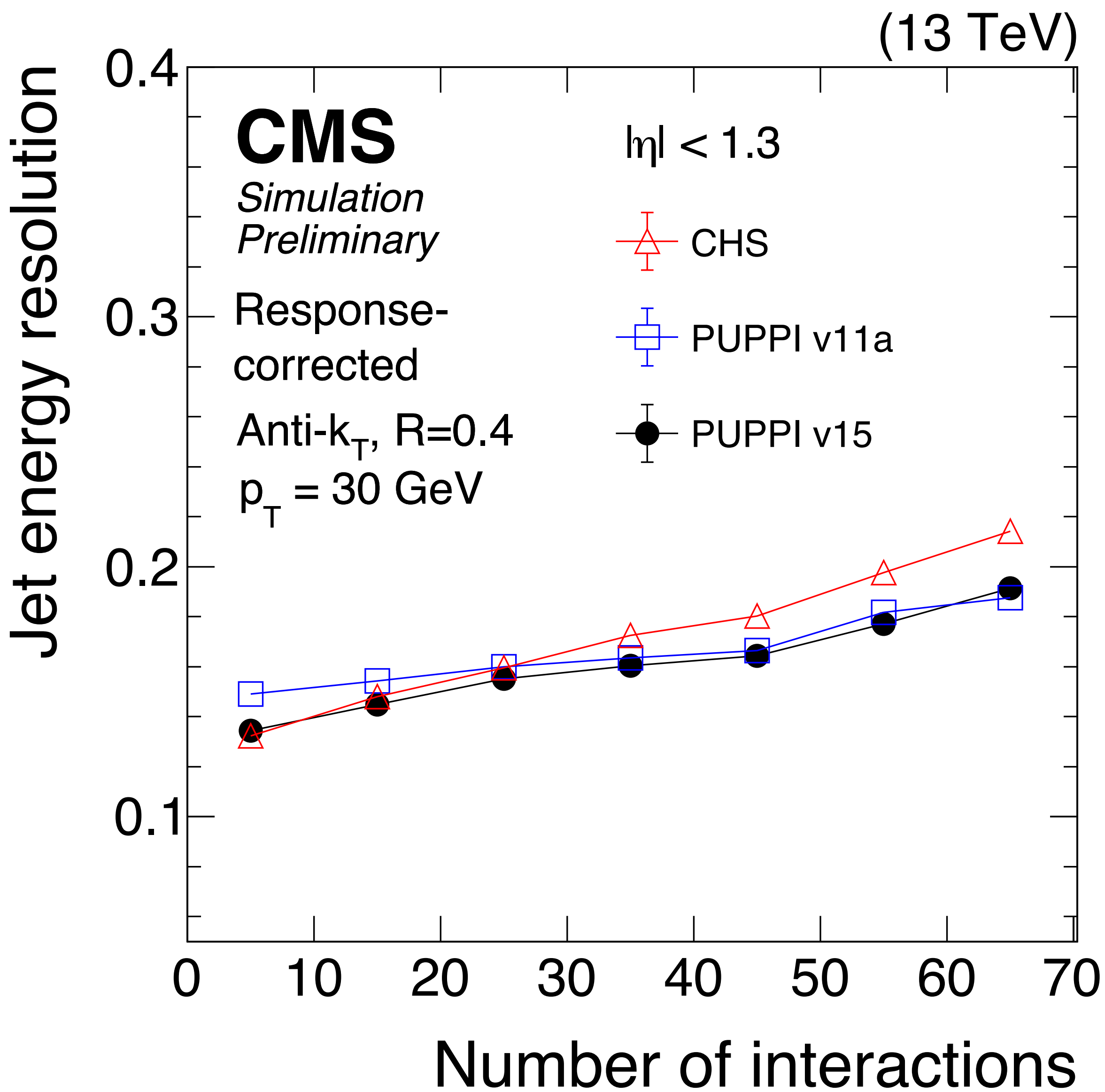
Noise Jet ID

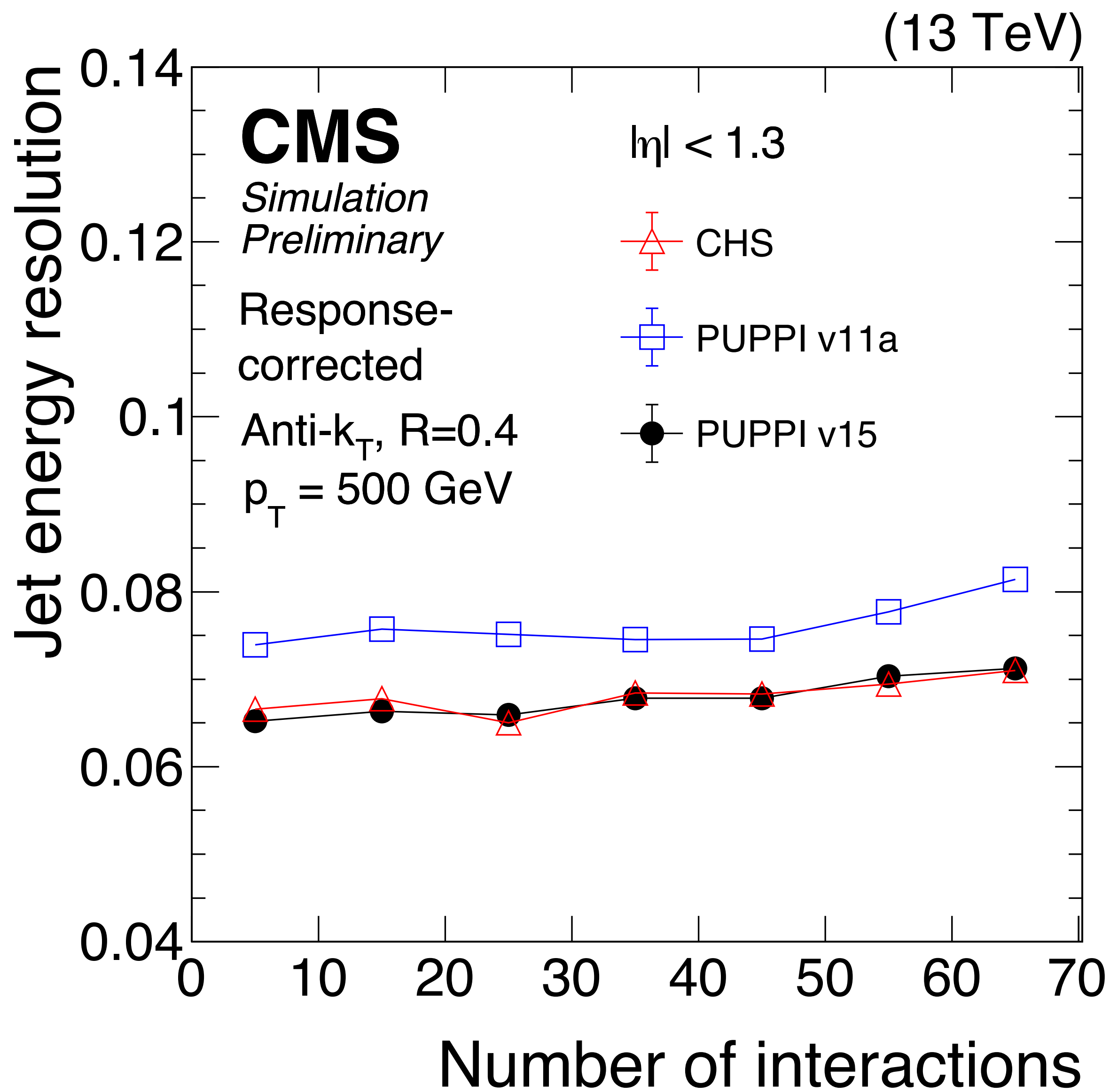




Purpose: remove noise from HCAL & ECAL by retaining ~98% of the real jets

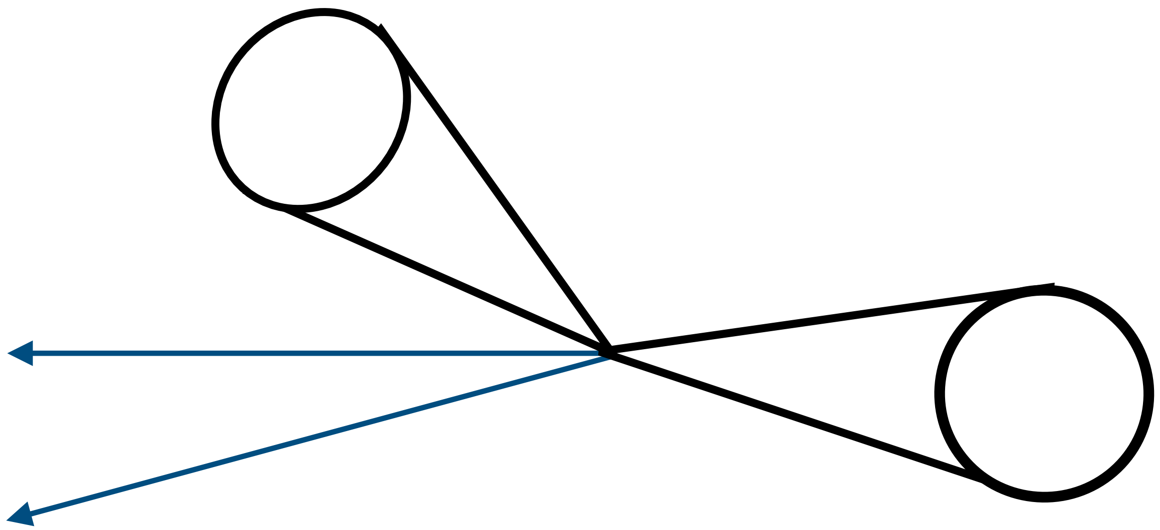
Jet energy resolution





Efficiency & purity in Z+jets

generator jets



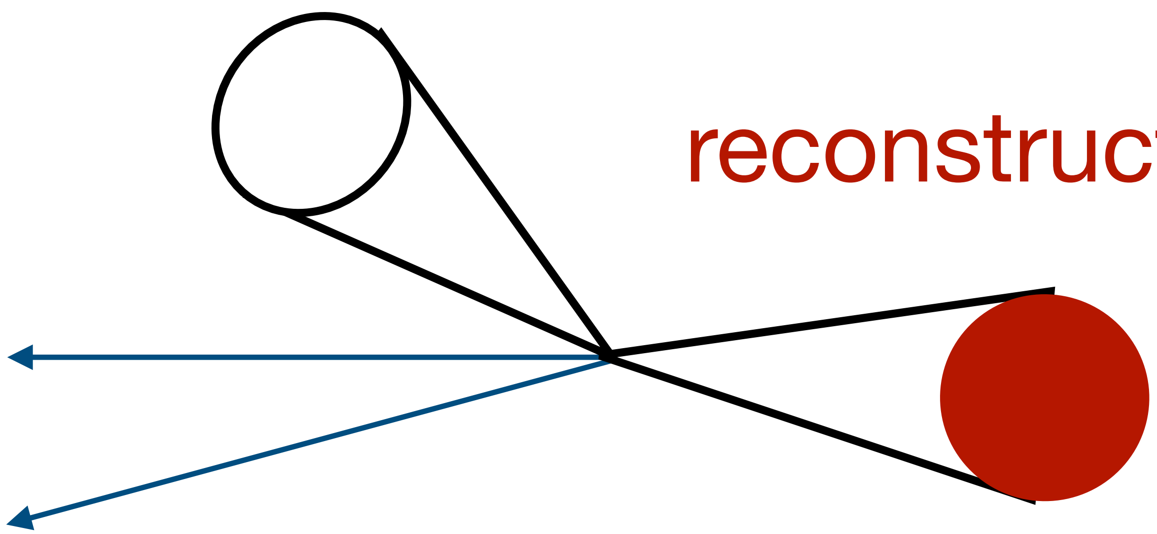
Efficiency

Simulation only study

Efficiency & purity in Z+jets

generator jets

reconstructed jets



Either not reconstructed or rejected:

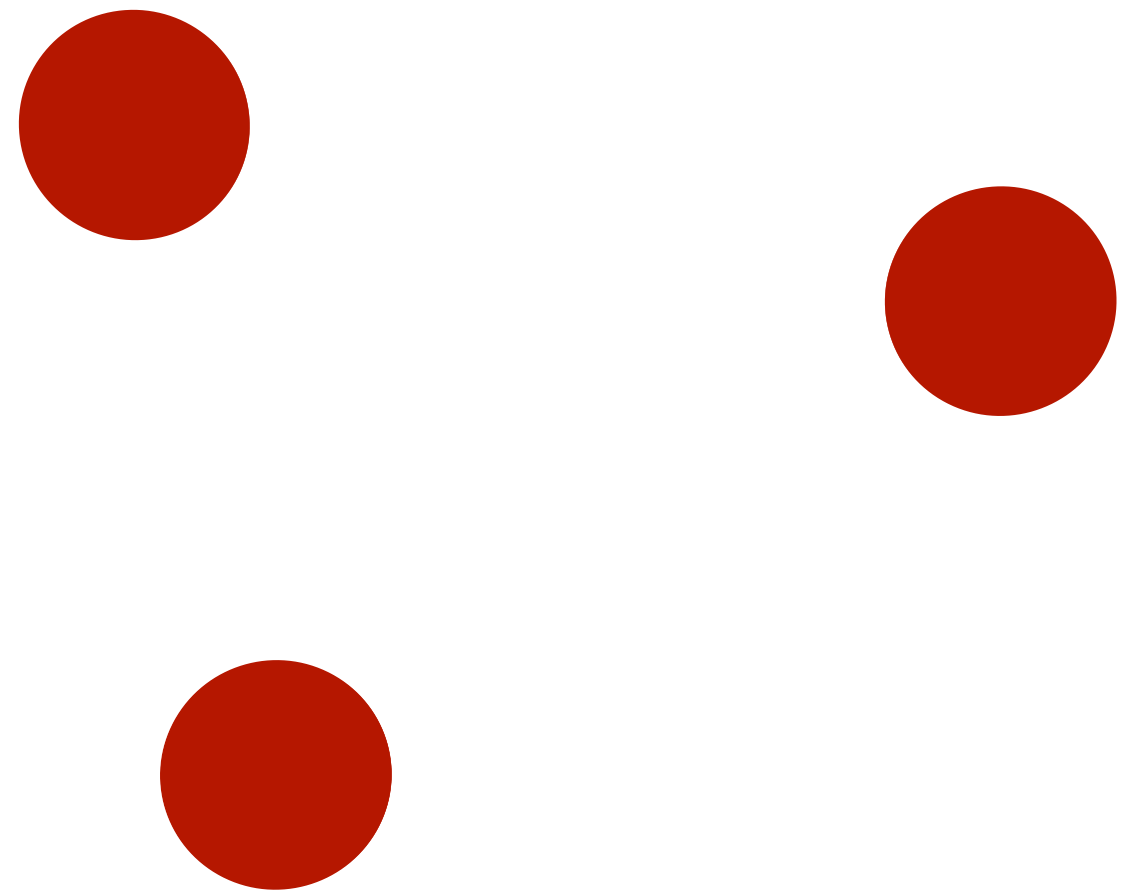
- Pileup Jet ID
- down weighted particles from PUPPI
- …

Efficiency

Simulation only study

Efficiency & purity in Z+jets

reconstructed jets



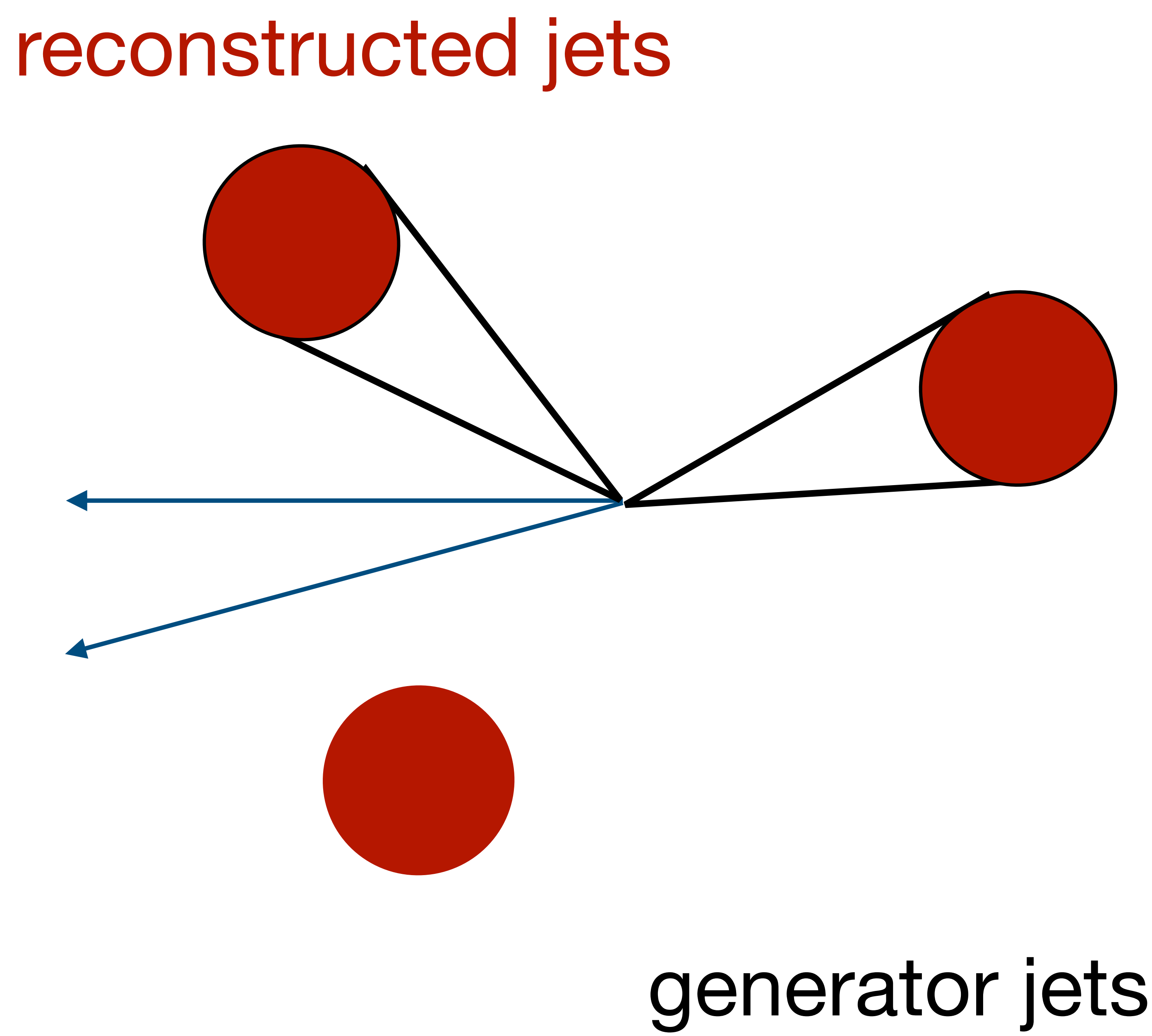
Purity

Simulation only study

Efficiency & purity in Z+jets

Most likely PU Jets did not get rejected:

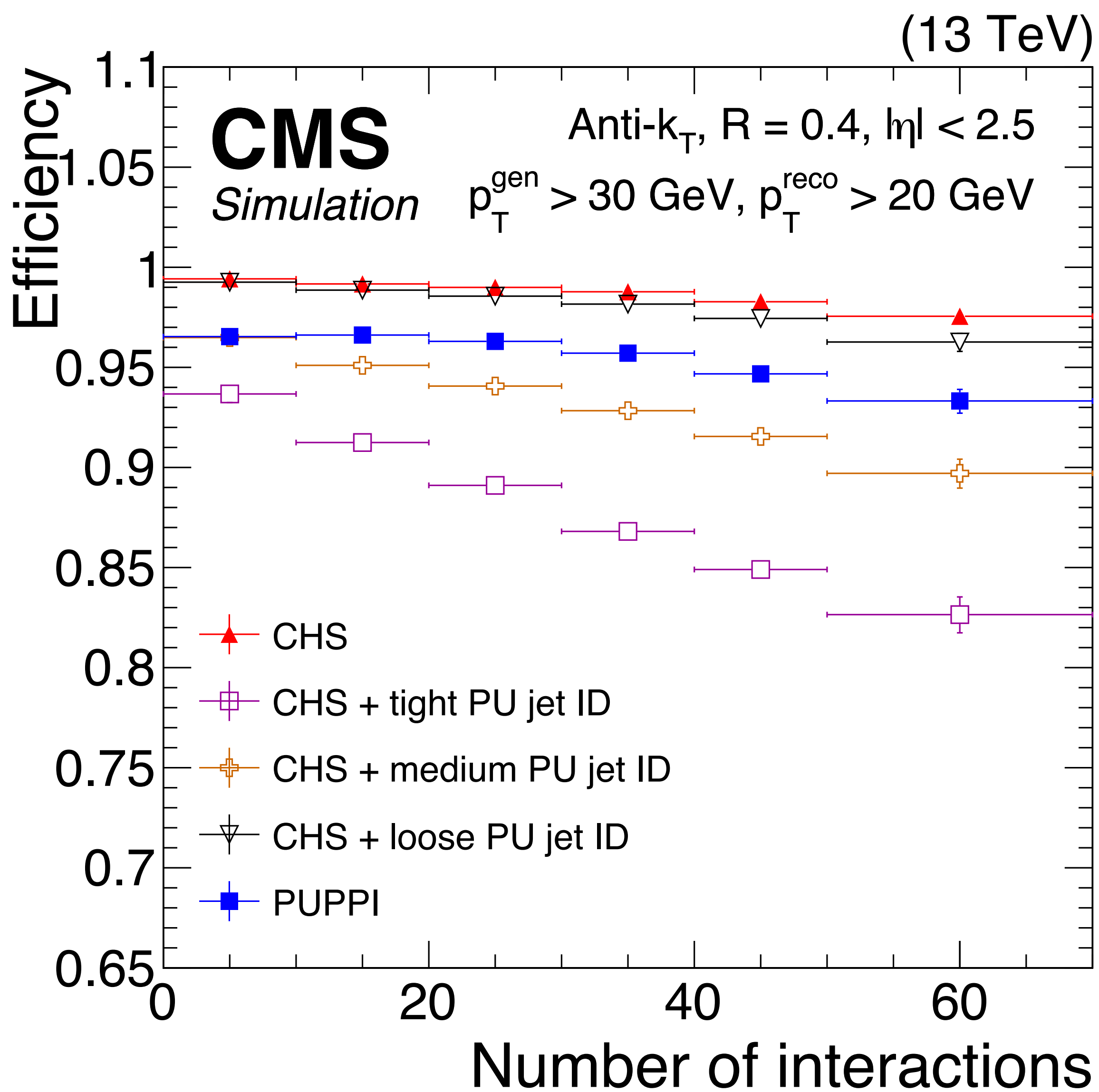
- CHS does not remove enough jets
- PUPPI does not remove enough particles
- …



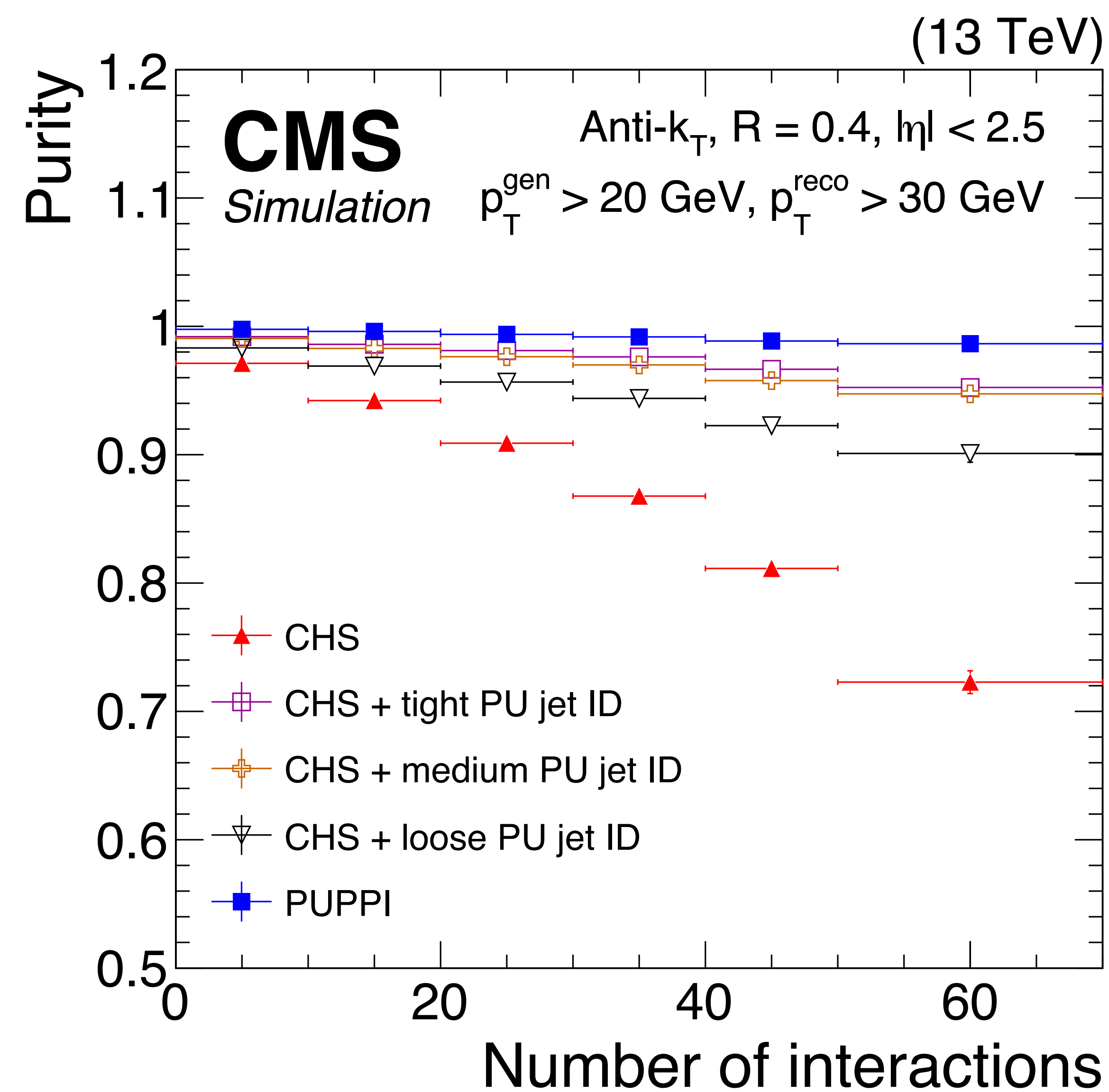
Purity

Simulation only study

Efficiency and purity

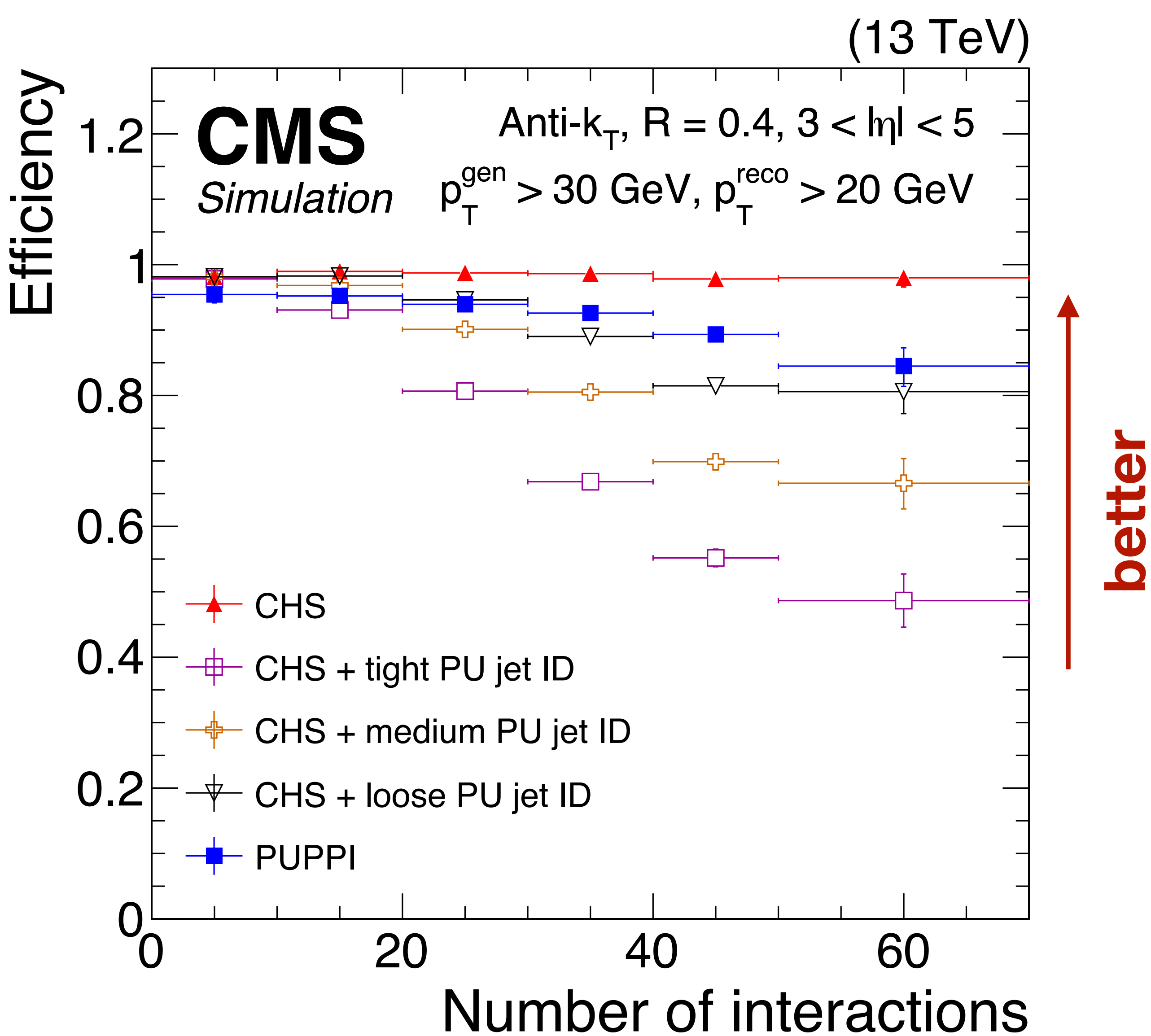


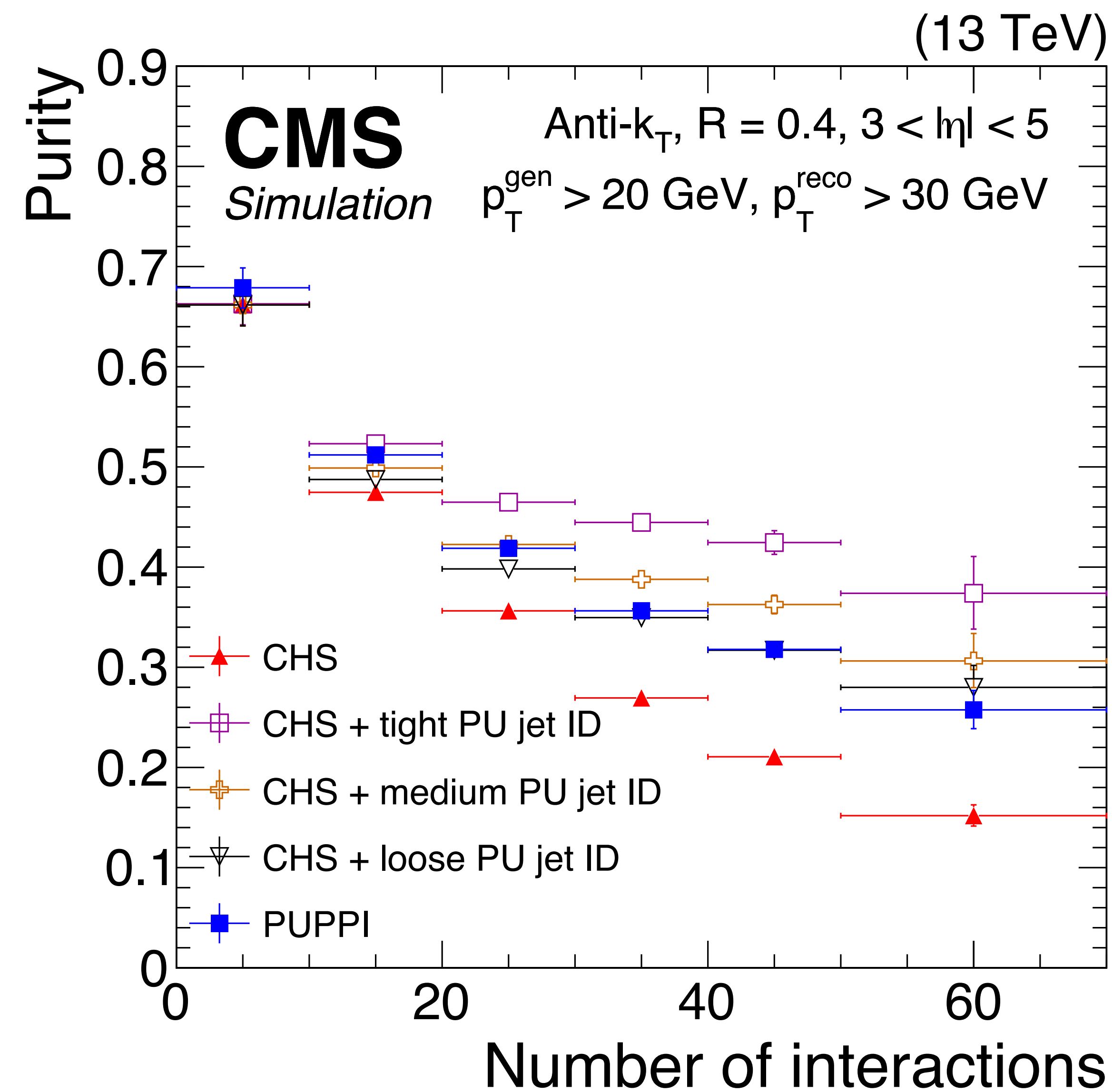
better



PUPPI better than **CHS+loose/medium** pileup jet ID

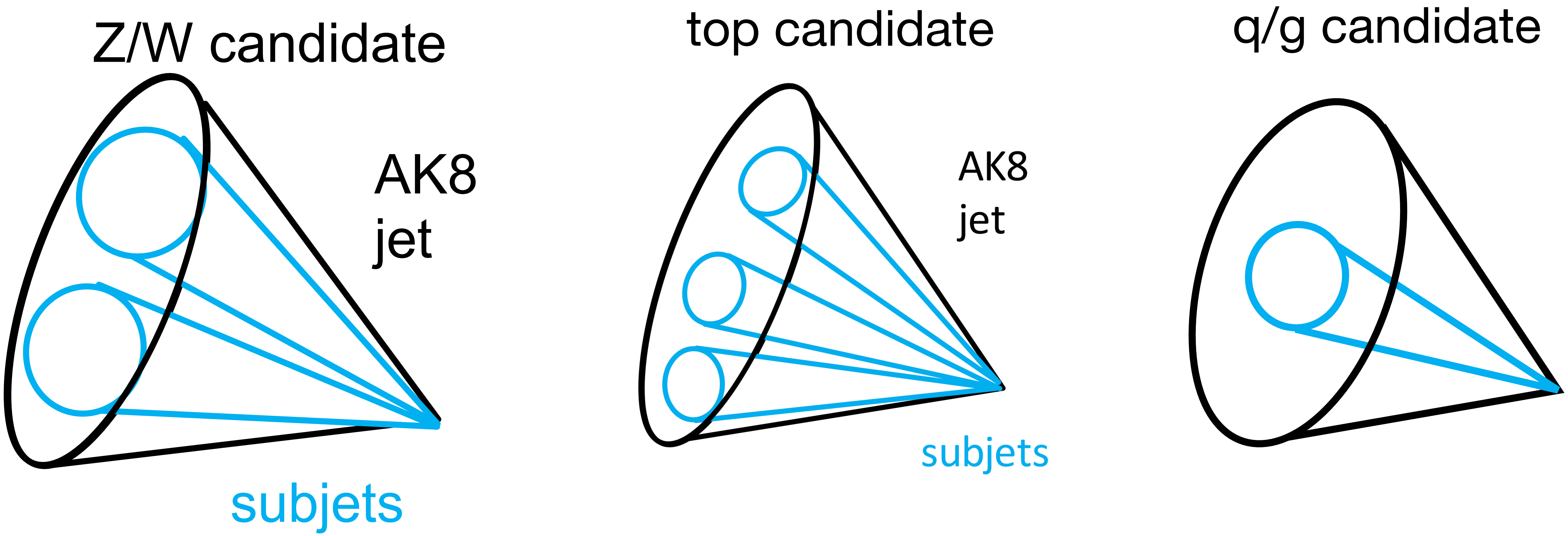
Efficiency and purity





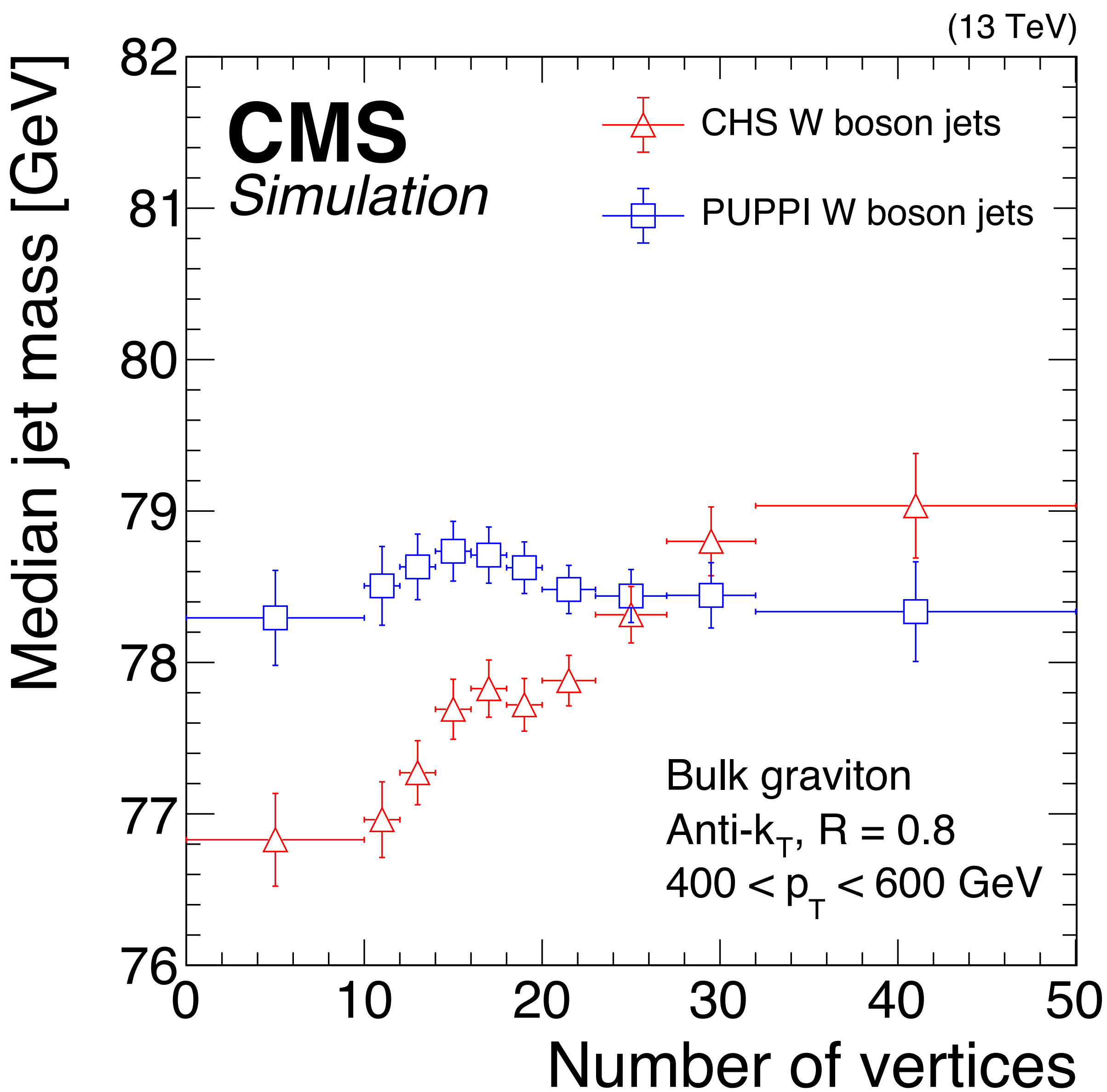
PUPPI better than **CHS+loose/medium** pileup jet ID

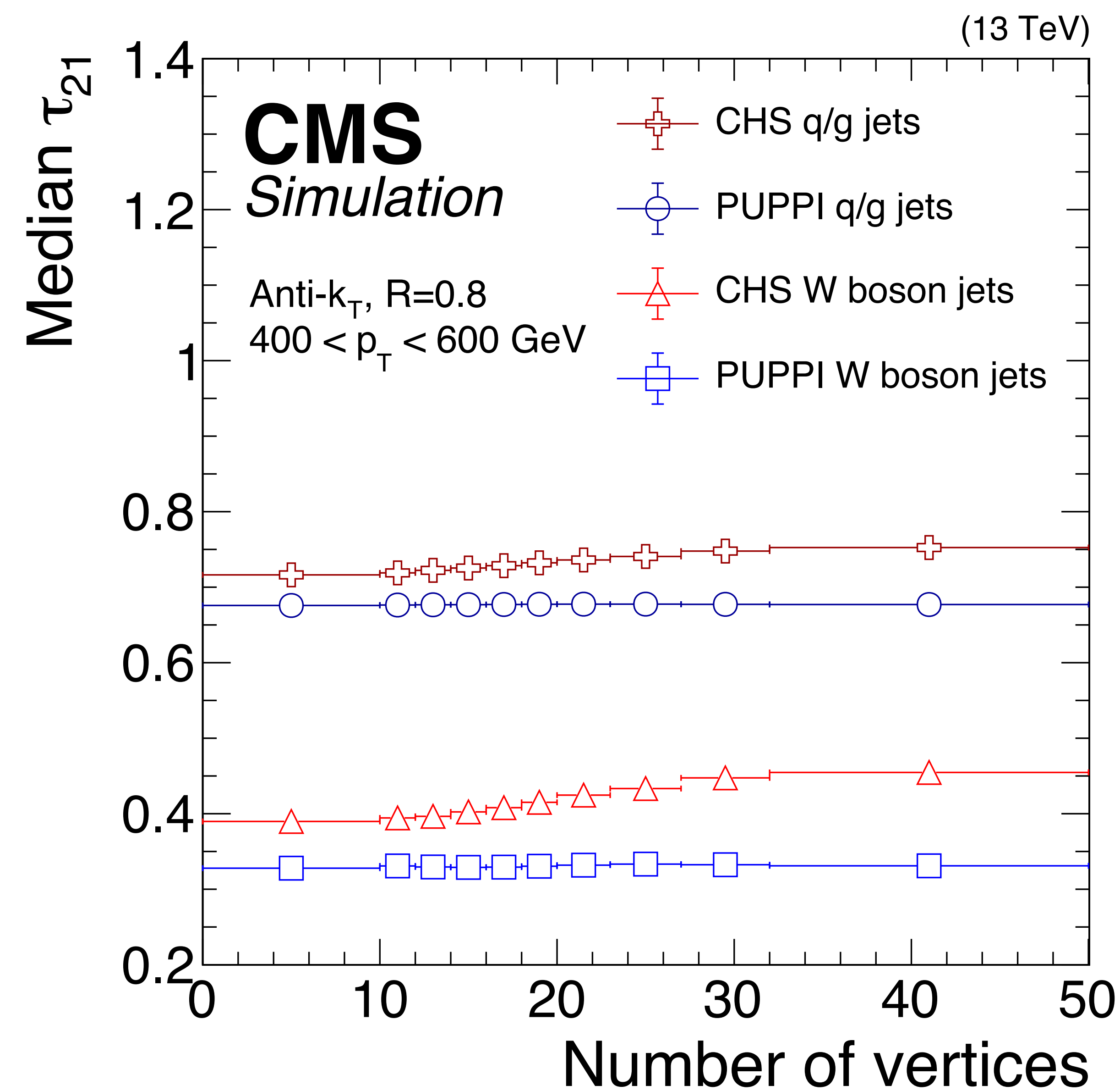
Identifying W bosons



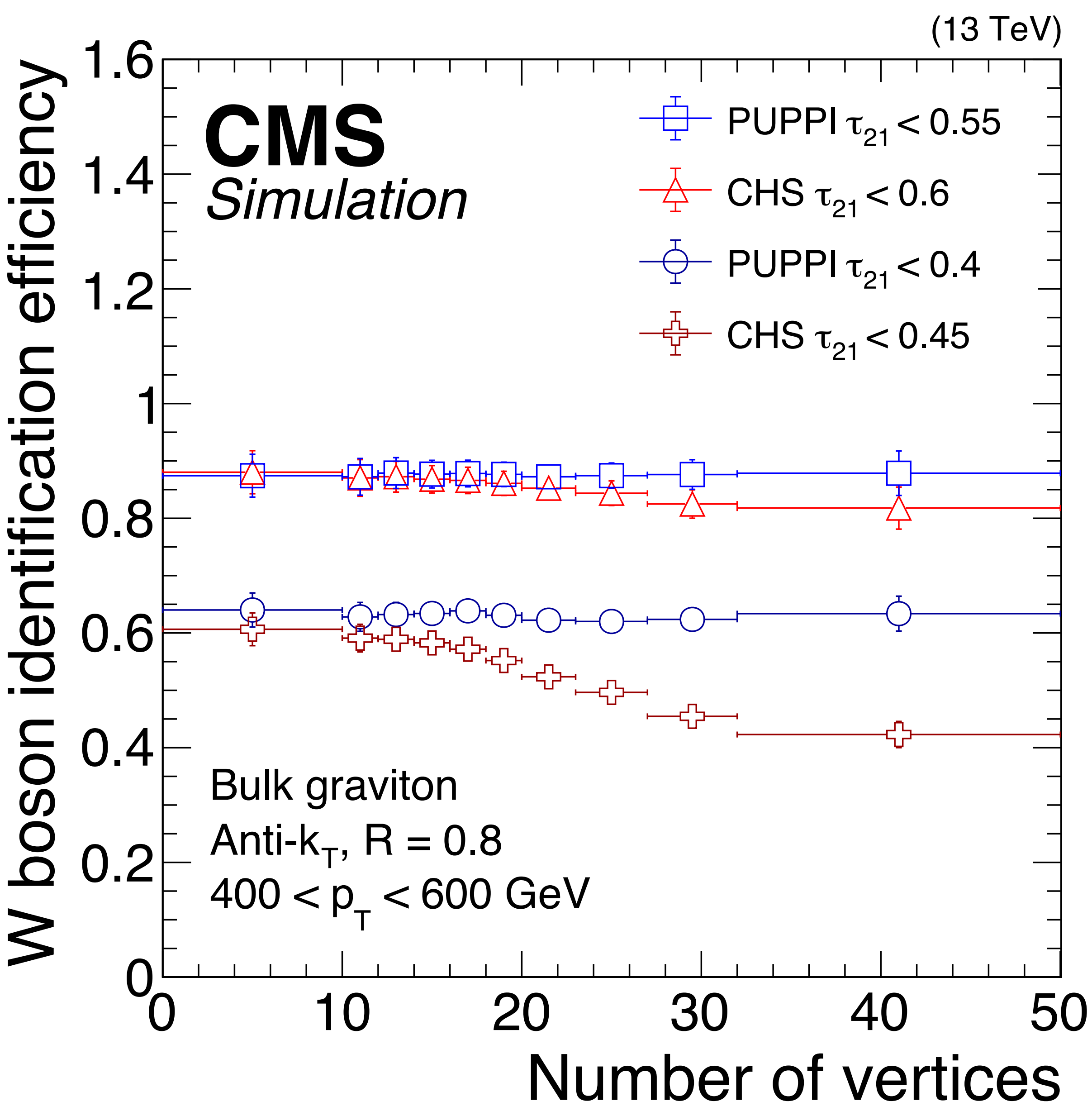
Identified through SD mass, n-subjettiness, subjet b-tagging, ...

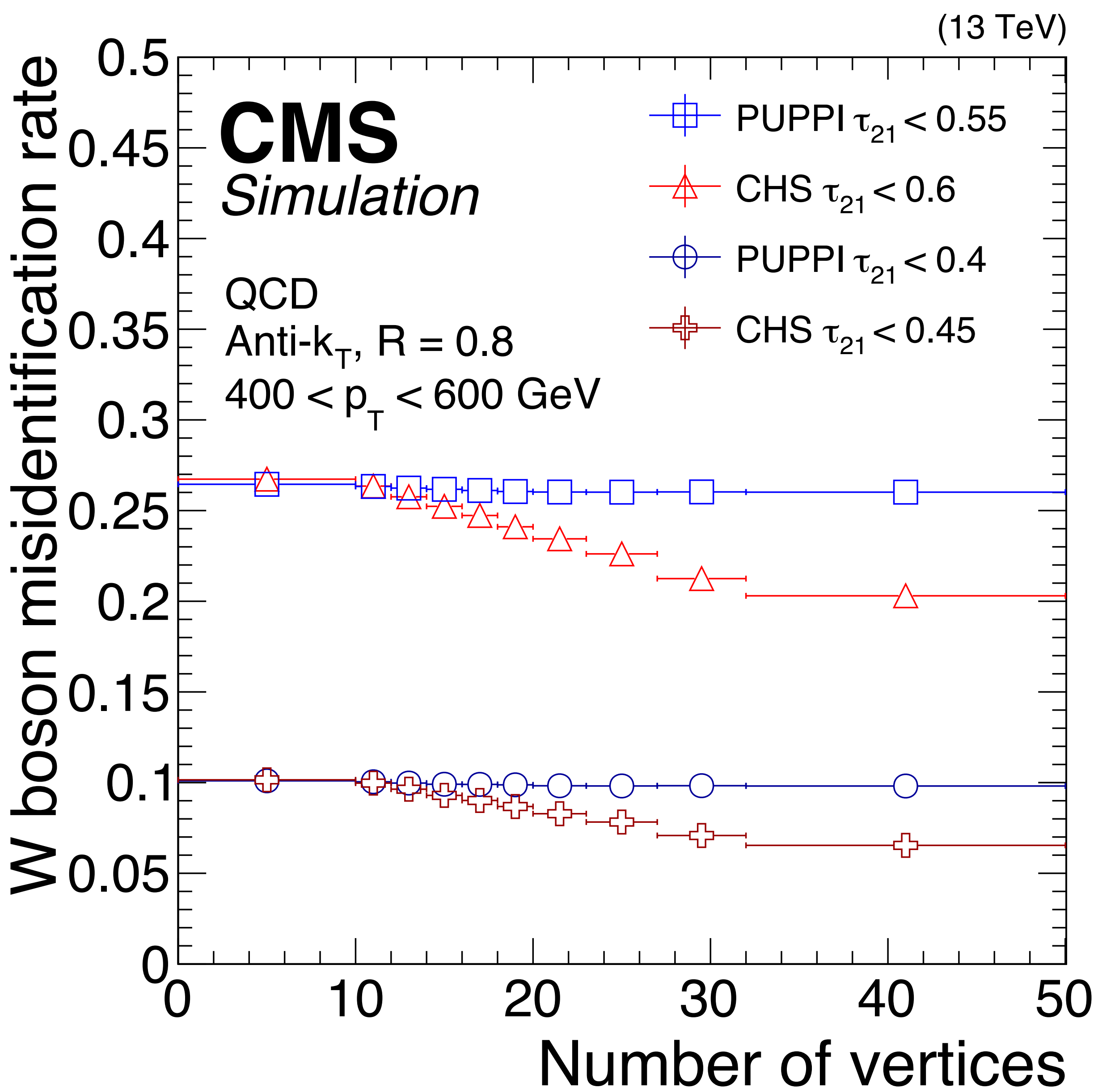
Identifying W bosons



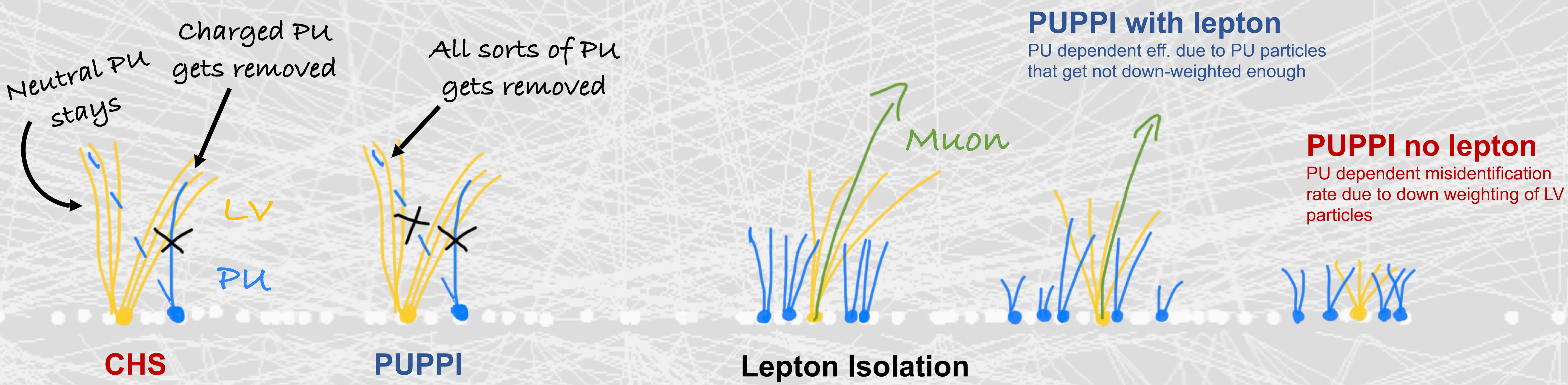


Identifying W bosons





Pileup Mitigation Techniques in CMS



- full validation of PUPPI shows a good performance, especially at high PU

  CMS-JME-18-001

- PUPPI is the standard pileup mitigation technique for CMS in Run 3

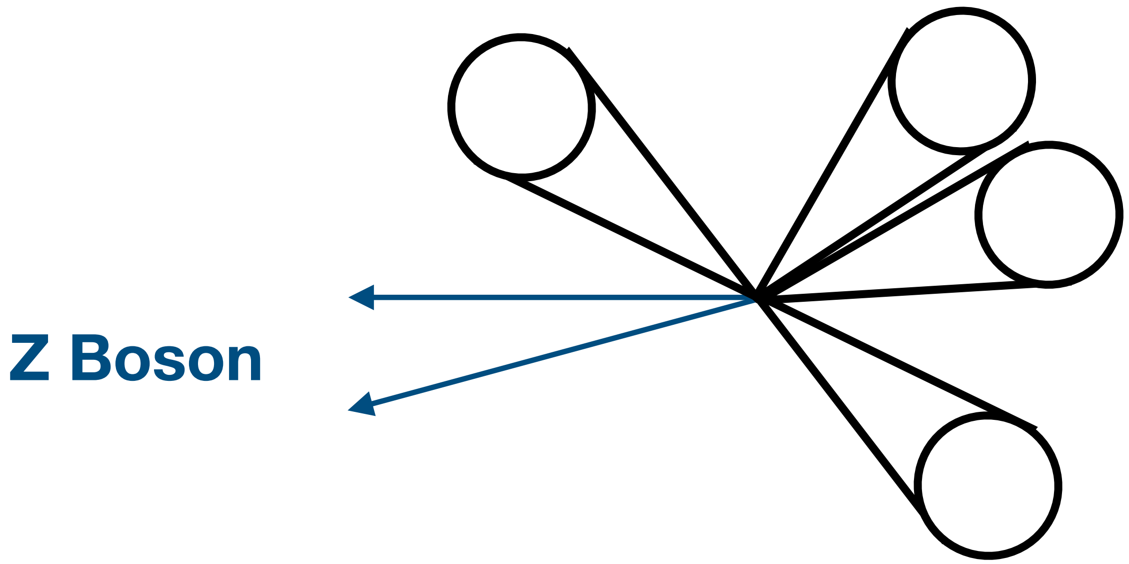
  CMS-DP-21-001

Backup

PU Jet ID variables

Input variable	Definition
LV $\sum p_T$ fraction	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_{ringX}, $X =$ 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 $\sqrt{\sum_i p_{\mathrm{T},i}^2} / \sum_i p_{\mathrm{T},i}$

Hadronic recoil performance

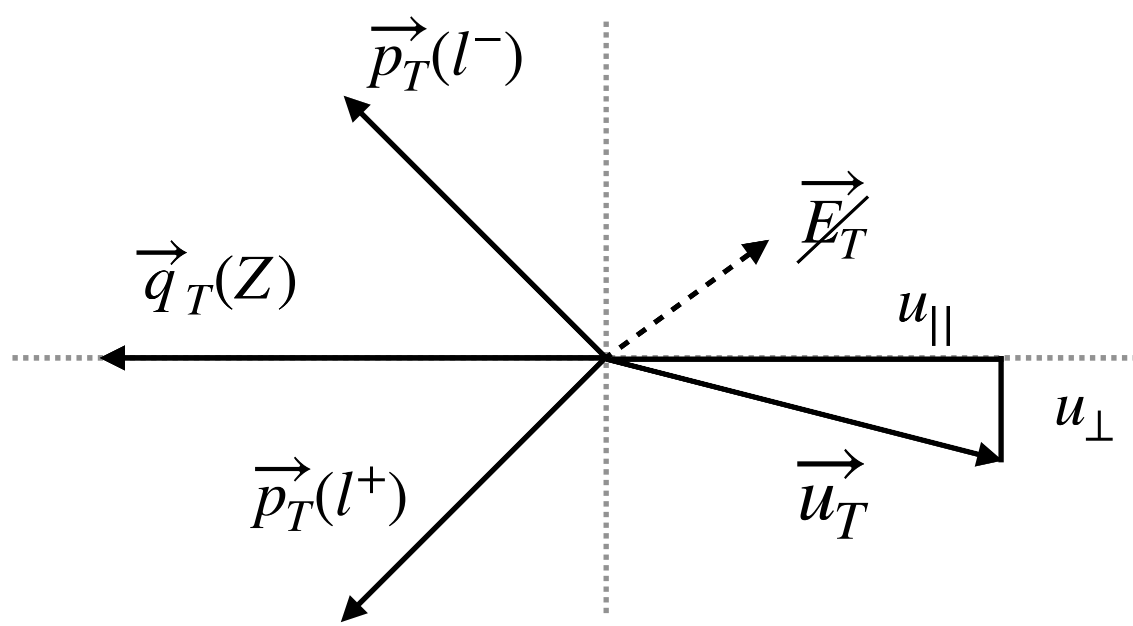


1. Sum up the AK4 jets in an event = hadronic recoil to the Z boson
2. In the ideal case:

$$p_T^Z - \sum p_T^{\text{jets}} = 0$$

Response:

$$\frac{u_{||}}{\vec{q}_T(Z)}$$

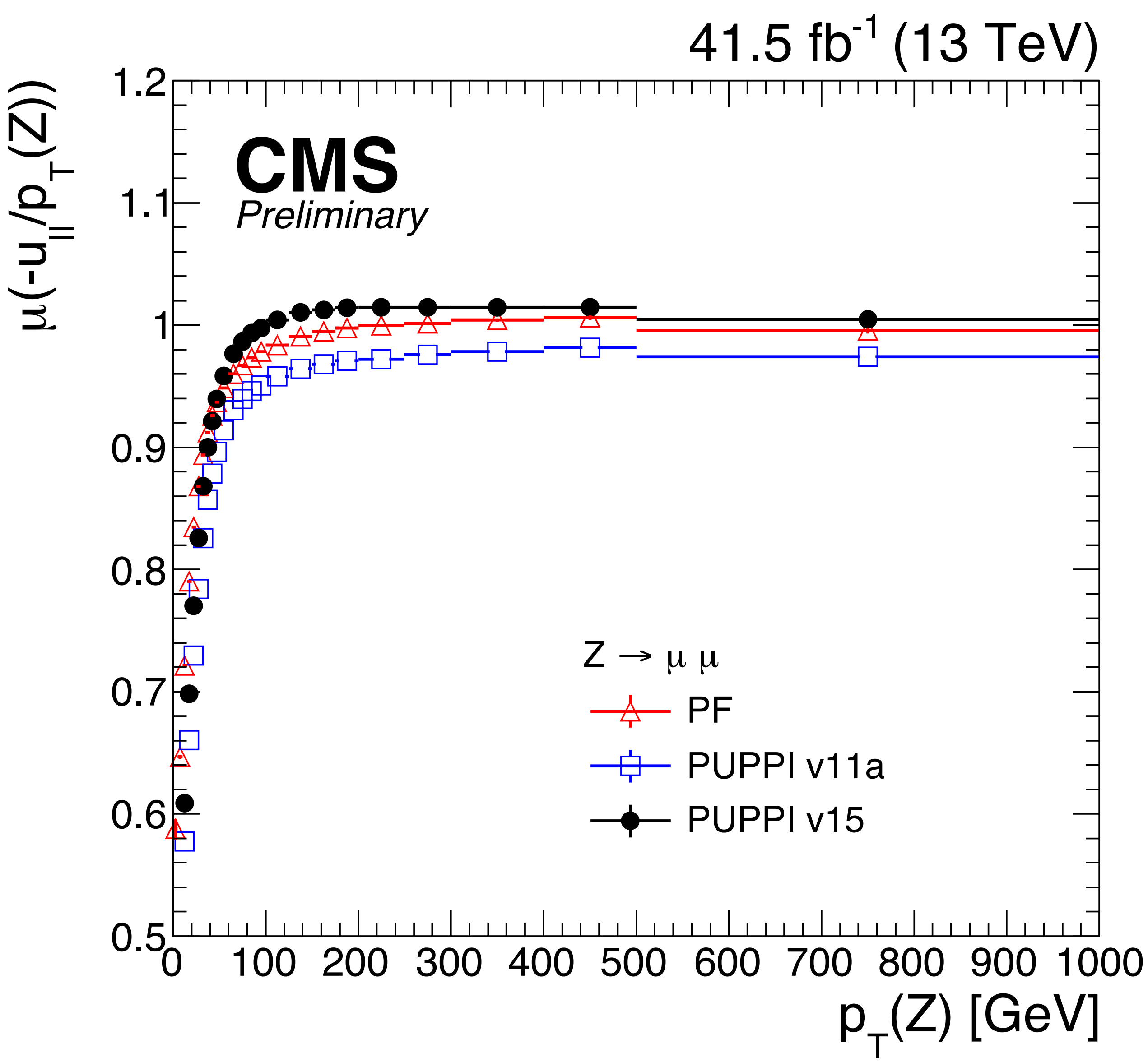


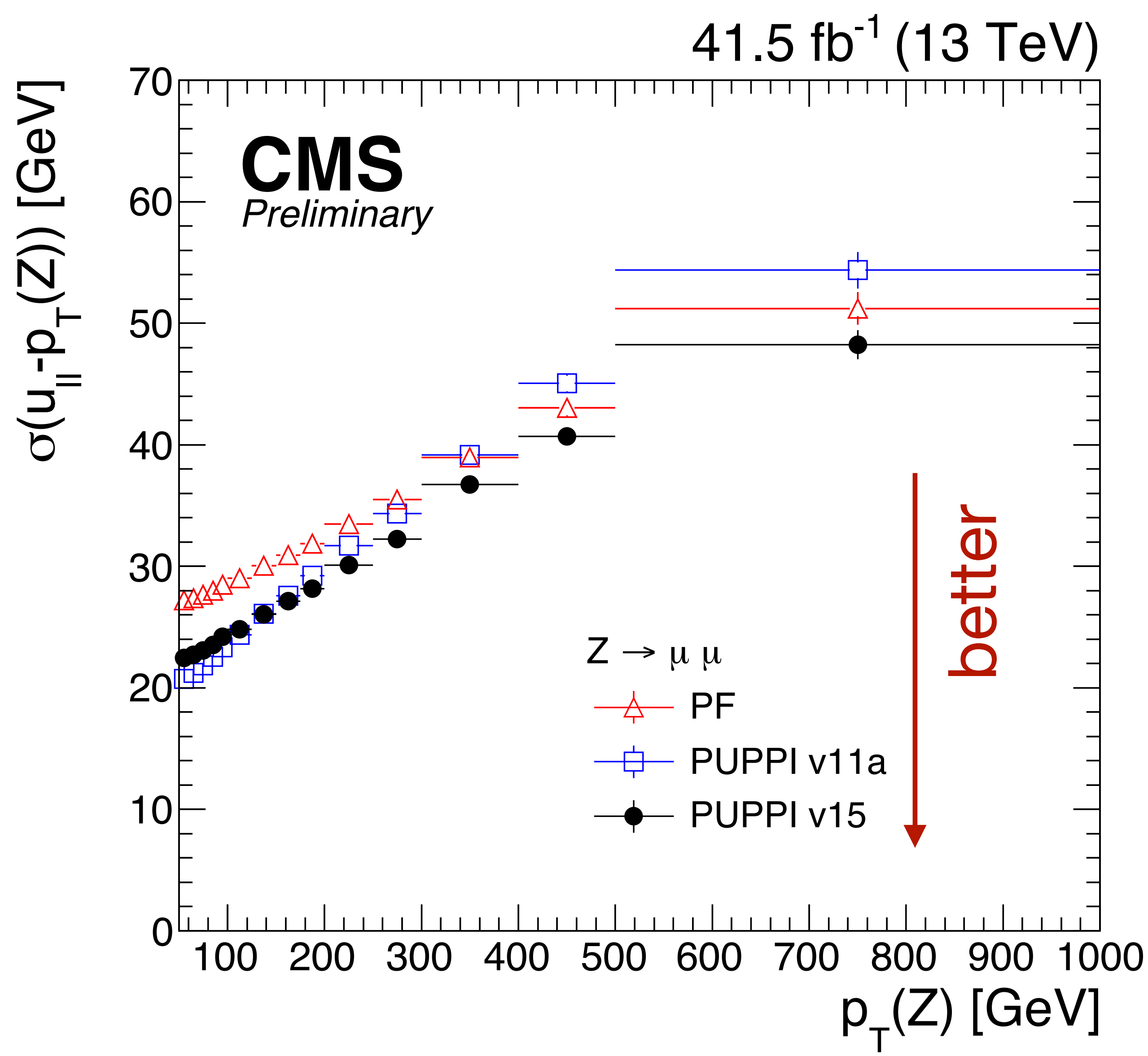
Resolution:

$$\sigma(u_{||})$$

$$\sigma(u_\perp)$$

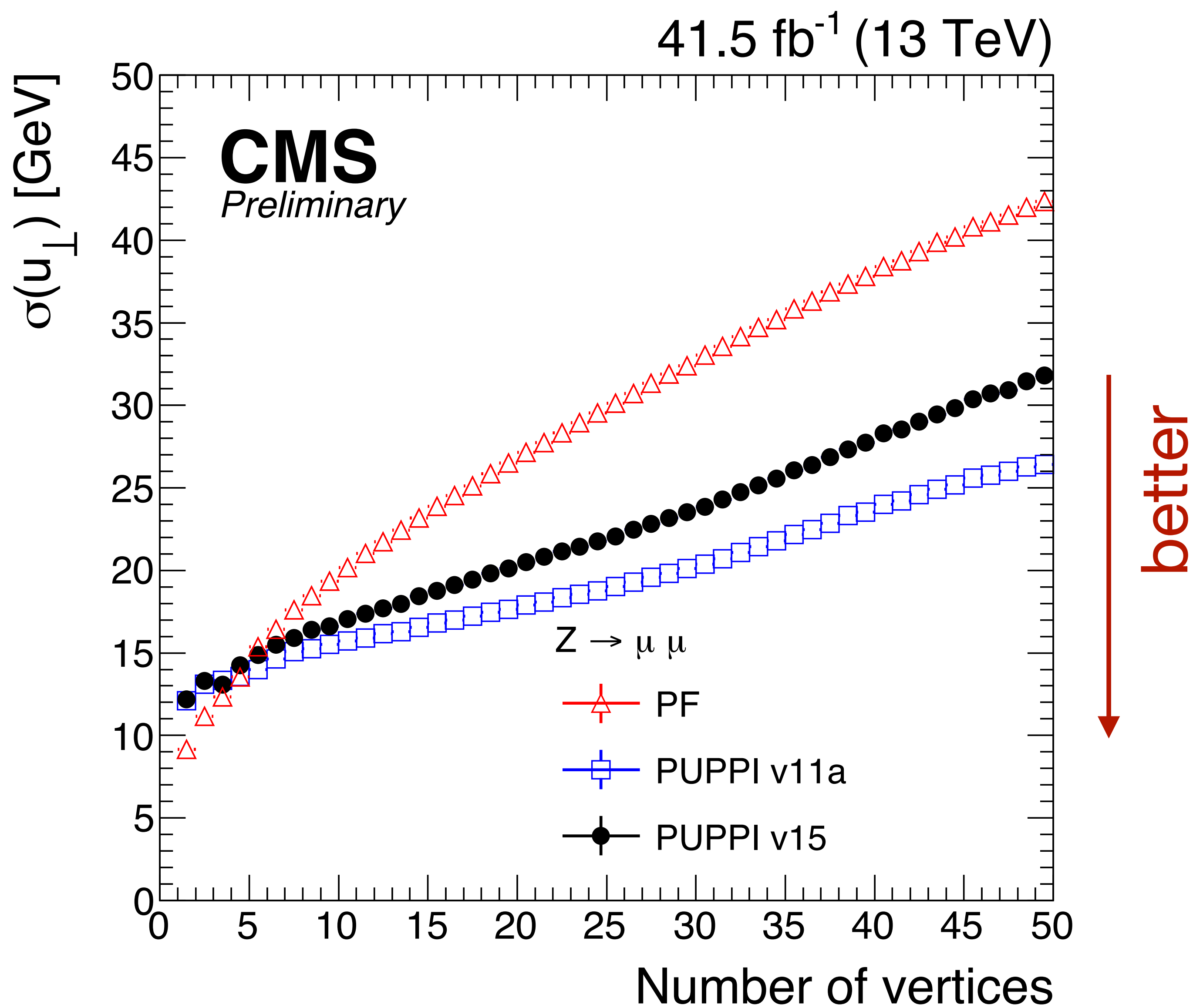
Hadronic recoil performance





Response (left) and resolution (right) both improved with **PUPPI v15** compared to **PUPPI v11a** and is better than **PF**

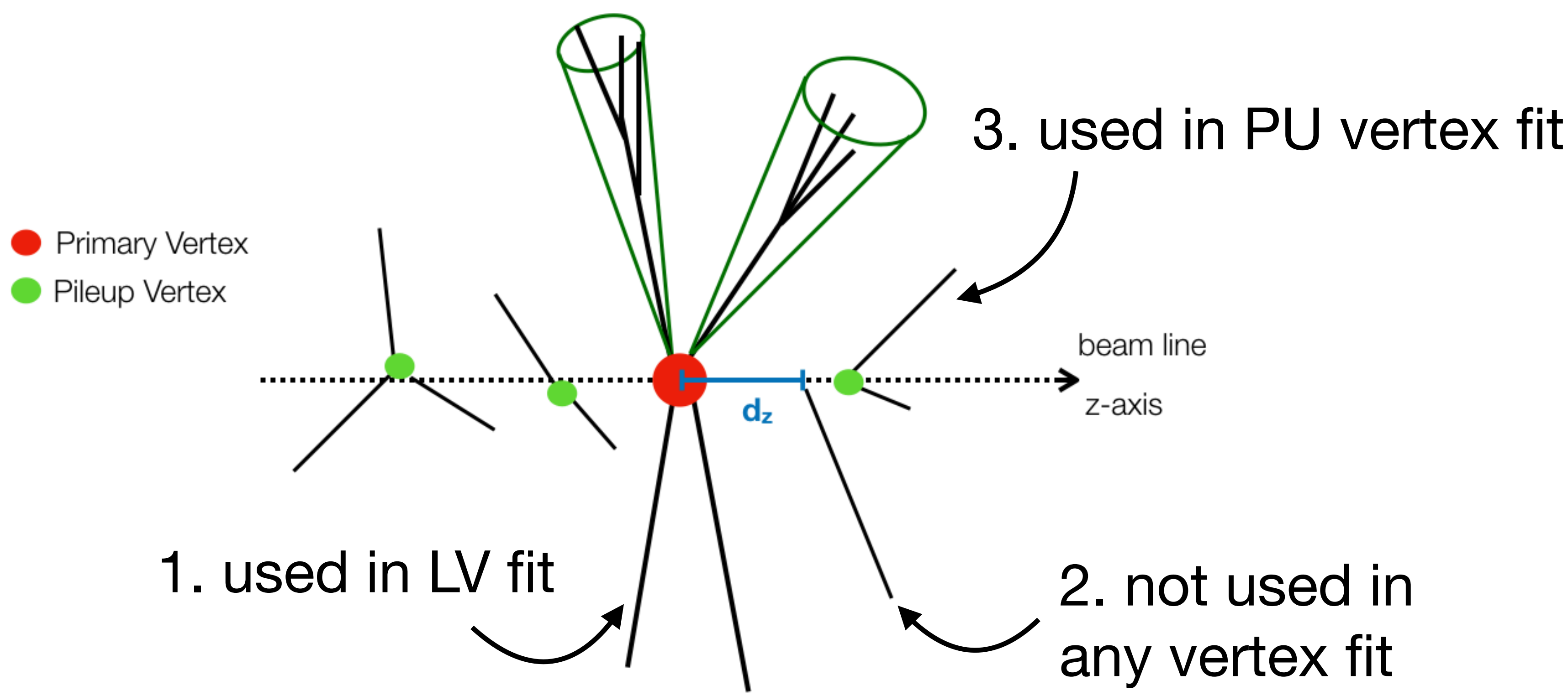
Hadronic recoil performance



Resolution of $u_{\perp}$ slightly worse with **PUPPI v15** compared to **PUPPI v11a** but still better than **PF**

PUPPI v15

Track-vertex association



	CHS	PUPPI v11	PUPPI v15
used in LV fit	keep	keep	keep
used in PU vertex fit	reject	reject	if 1st or 2nd PU vertex && $\|d_z\| < 0.2$ cm keep else reject
not used in any vertex fit	keep	keep if $d_Z < 0.3$ cm	*η-dependent see next slide

PUPPI v15

	$p_T > 20$ GeV	$p_T < 20$ GeV
$\|\eta\| < 2.4$	keep	calculate a weight
$\|\eta\| > 2.4$	keep	if $\|d_Z\|$ < 0.3 cm keep else reject

Tab. 2: Categories for charged particles that are not used in the fit of the LV or a PU vertex.

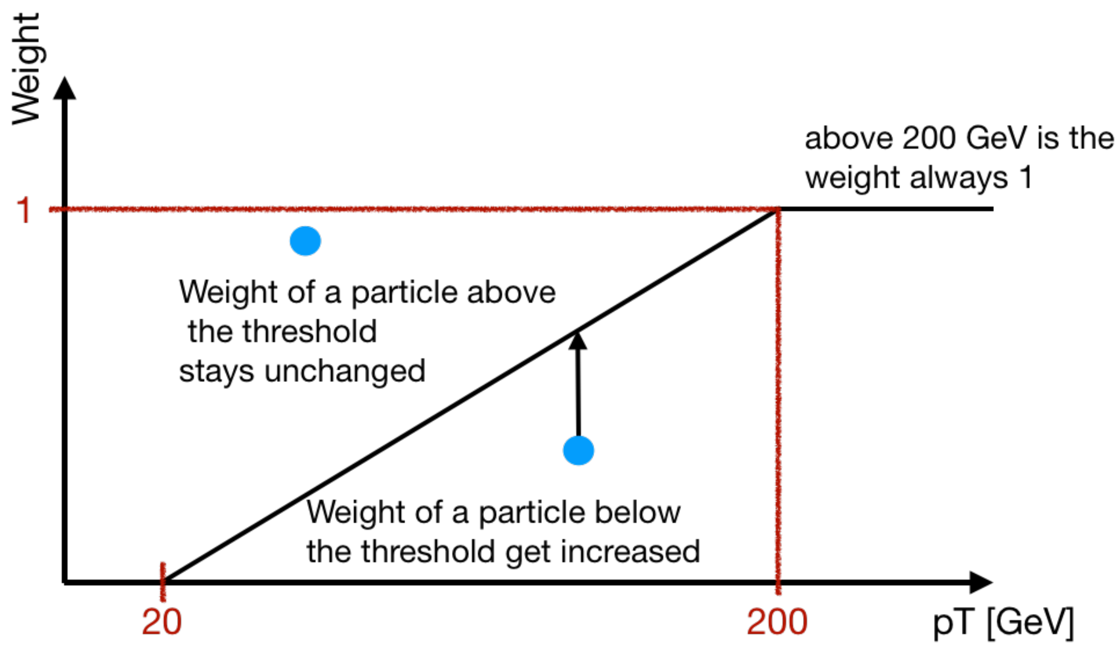
If a charged particle is not used in any fit PUPPI keeps the charged particle only if $|d_Z|$ < 0.3 cm, while CHS keeps all charged particles. However, the track-vertex assignment, especially $|d_Z|$ is not working for high-p_T particles and therefore rejects too many LV particles. In the new tune we therefore implemented a protection for this.

In $|\eta| < 2.4$: Charge particles with p_T>20 GeV are kept, while the ones with p_T<20 GeV are treated as neutral. These particles get a weight and are excluded from the α calculation in $|\eta| < 2.5$ (see Tab. 2).

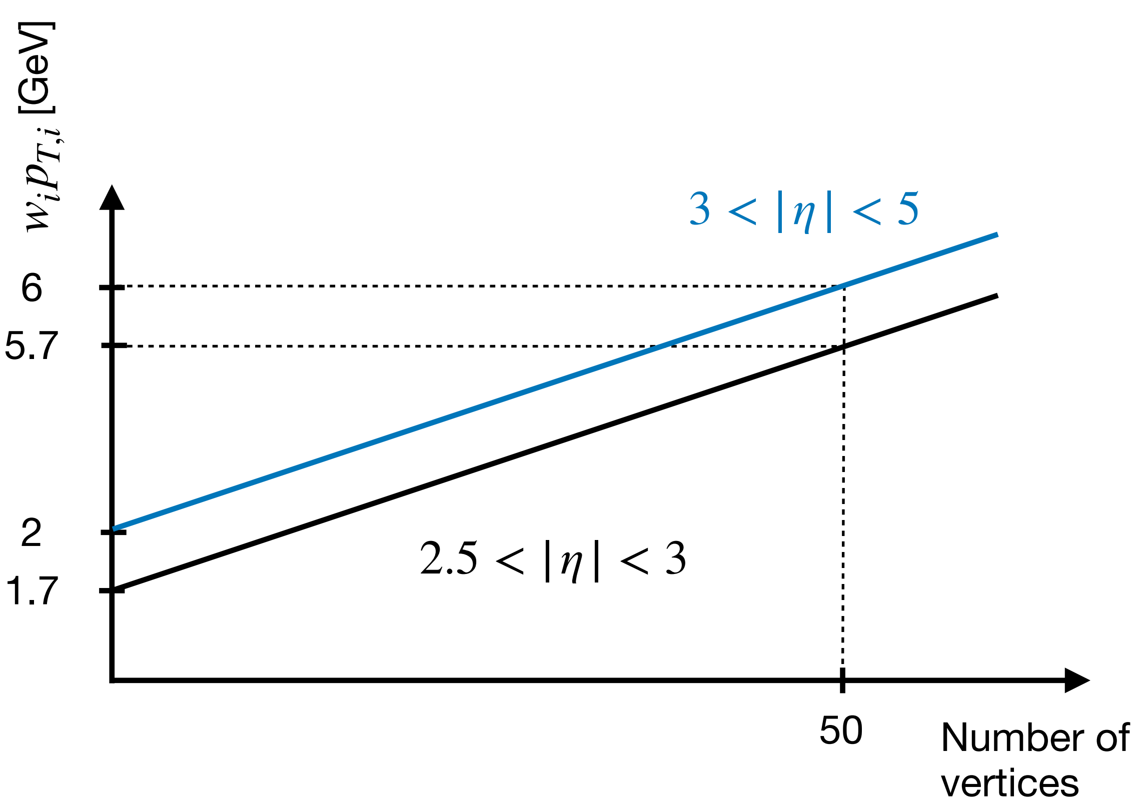
In $|\eta| > 2.4$: Charged particles with $p_T > 20$ GeV are kept. Charged particles with a $p_T < 20$ GeV are kept if $|d_Z|$ < 0.3 cm (see Tab. 2).

The motivation behind the $|\eta| < 2.4$ condition comes from the fact that only tracks satisfying this condition are used for vertex fitting. PUPPI internally is split into $|\eta| < 2.5$ and $|\eta| > 2.5$ since $|\eta| = 2.5$ is the boundary of our Phase-0 tracking system. The tracking system was extended to $|\eta|<3.0$ with the Phase-1 upgrade.

PUPPI v15



In addition, neutral particles receive a protection based on their p_T. If the weight of a neutral particle is less than $w_i < p_T \cdot \frac{1}{200-20} - \frac{20}{200-20}$ (black line shown in the bottom left plot) than the weight gets increased to be equal.



The weight is multiplied to the four-momentum of the corresponding particle. Each particle has to pass a certain p_T threshold to be considered in the clustering. The p_T threshold varies between 1.7 GeV and 6 GeV depending on the number of vertices in the event and the η of the particle (plot on the bottom left).

Efficiency and purity

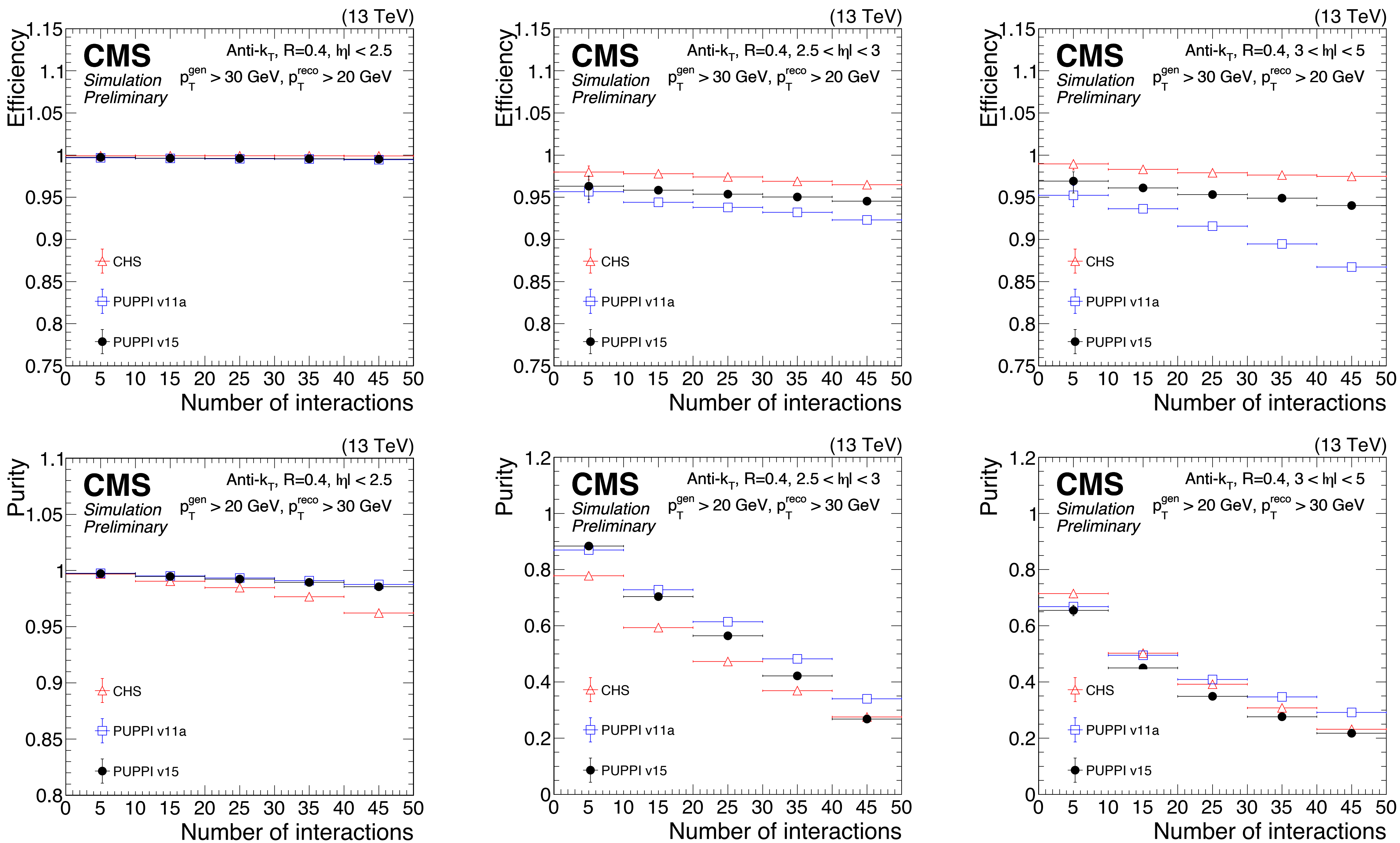


Figure 4: The LV jet efficiency (upper) and purity (lower) in Z+jets simulation as a function of the number of interactions for PUPPI v11a (blue open squares), CHS (red open triangles), and the new tune of PUPPI v15 (black filled circles). Plots are shown for AK4 jets with $p_T > 20$ GeV, and with (left) $|\eta| < 2.5$, (central) $2.5 < |\eta| < 3$ and (right) $|\eta| > 3$.