Performance of Jet Reconstruction at CMS

On behalf of the CMS collaboration
Christian Sander (University of Hamburg)

27th May 08 – HERA and the LHC workshop

Outline

- Overview of implemented jet algorithms
  - Iterative Cone
  - Midpoint Cone
  - SIS Cone

  Cone based algorithms

  - k_T

  Clustering algorithm
- The CMS calorimeter
- Calorimeter jet calibration
- Comparison of different jet algorithms
  - Reconstruction efficiency
  - η, ϕ and p_T resolution ...
- Performance in $t\bar{t}$ events
- Summary and plans

Jet Algorithms

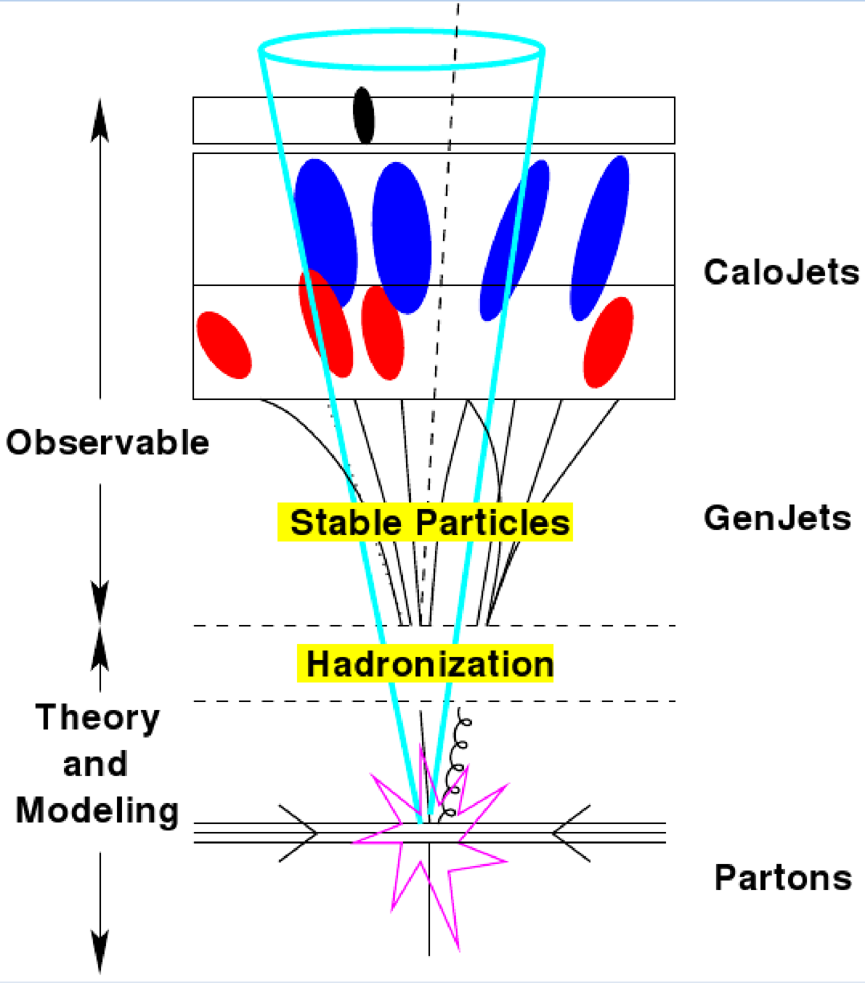
- Jet algorithm should provide good correspondence (multiplicity, position, energy) between

  calorimeter jet ↔ particle jet / parton

- Jet algorithms should be:

  collinear safe, infrared safe, fast ...

- Challenges:
  - Optimization of thresholds and zero suppression parameters at calorimeter tower and cluster level
  - Spurious contributions from noise and multiple interactions (pileup)
  - Handling of Underlying Event
  - Response calibration
  - Out-of-cone showering effects



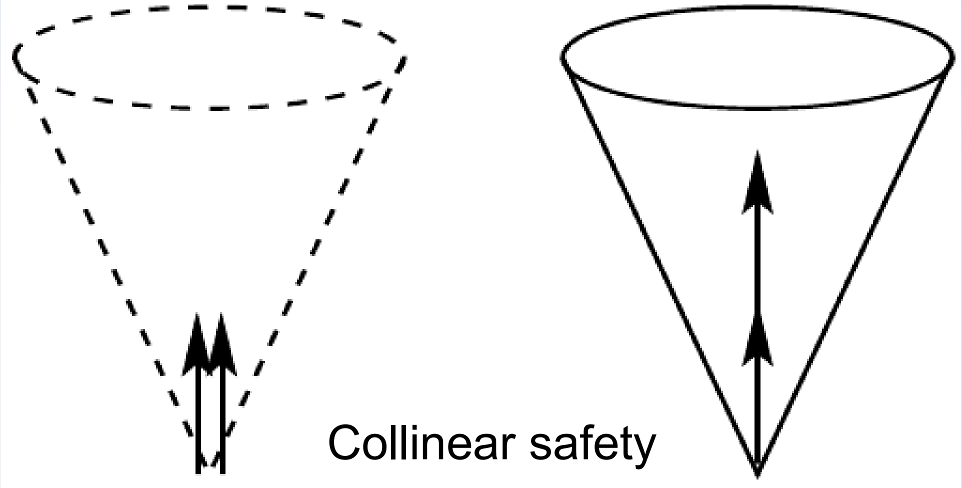
Iterative Cone

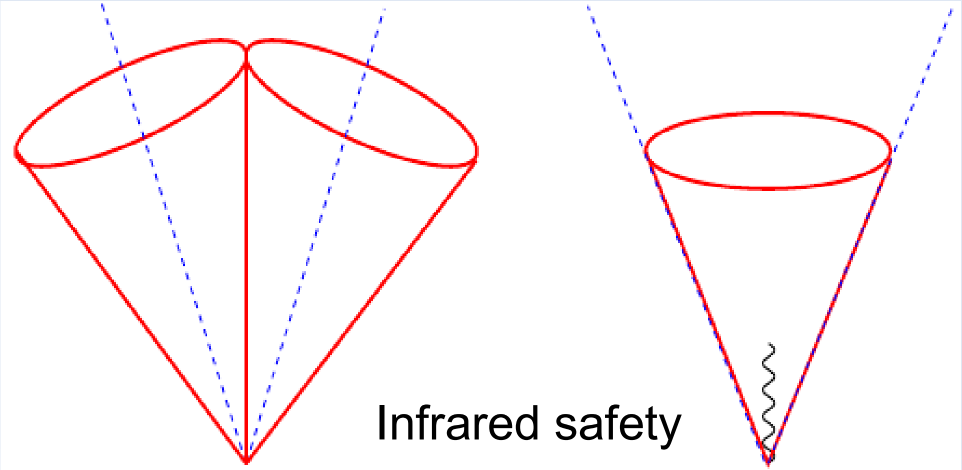
1) Start with p_T ordered list of objects
2) Choose first object as seed
3) Collect objects within a cone of radius R around seed
4) Recalculate jet axis and use it as new seed
5) Repeat from 3) until stable axis
6) Declare constituents as a jet and remove them from the input list
7) Repeat from 2) until list empty

Apply threshold on input list objects (save computing time and reduce noise)

→ Algorithm is neither infrared nor collinear safe

Fast and simple cone based algorithm
→ used by online / HLT



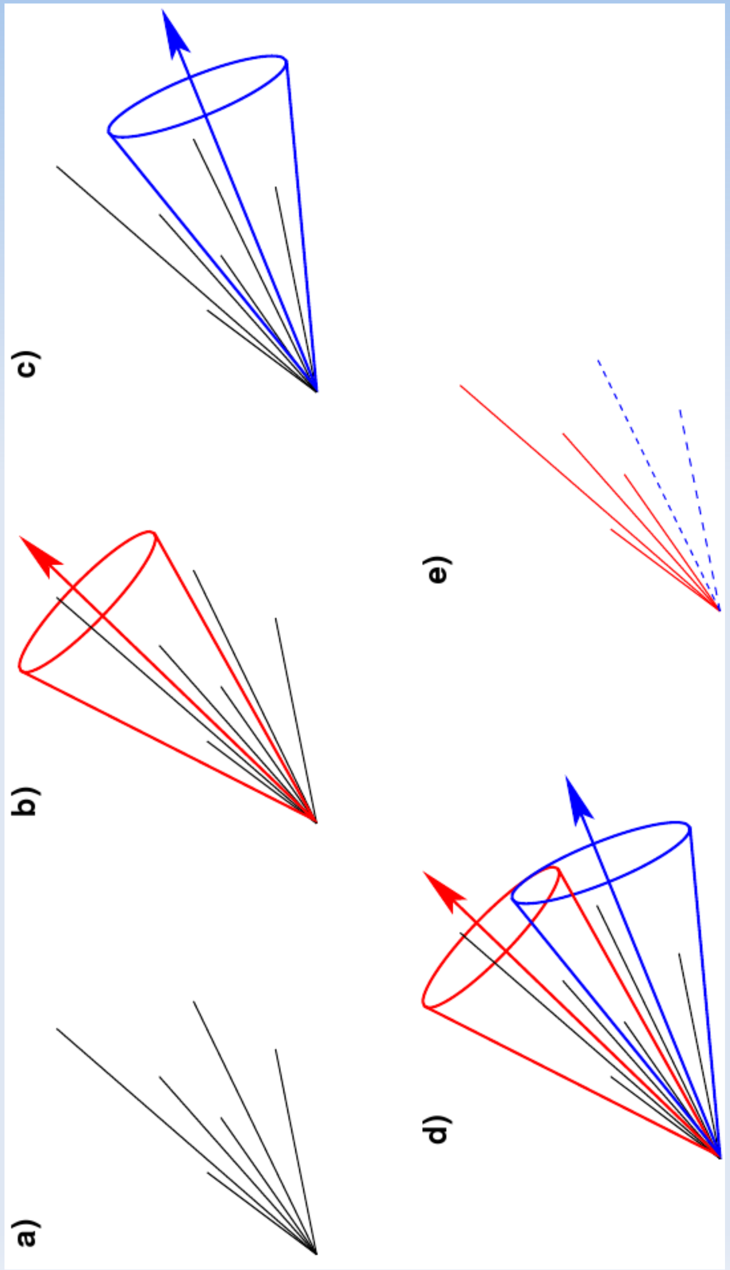


Midpoint Cone

- Similar to iterative cone, but:
  - Objects assigned to proto-jets are not removed from input list (overlapping proto-jets possible)
  - For each pair of proto-jets closer than R the midpoint is used as additional seed
  - Overlapping energy of 2 proto-jets is larger than 50% of smaller one

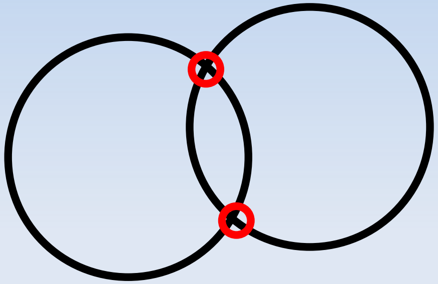
    → merge proto-jets

    → **Else:** split overlapping constituents according to distance to jet axis
- Midpoint Cone is not infrared safe beyond NLO
- Midpoint Cone is not collinear safe



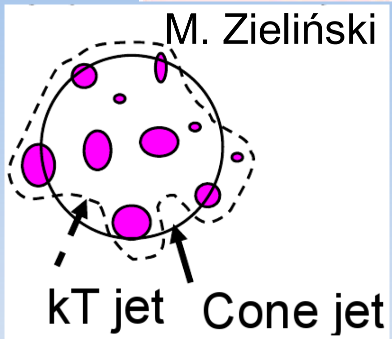
SIS Cone

- Seedless Infrared-Safe Cone Algorithm
  - **Test all possible cones:** for each two objects within $2R$ (defined by y rather than η) check the two cones with both objects on the circumference for stability



  - Splitting and merging similar to Midpoint Cone (but: scalar sum p_T ordering)
  - Some clusters might not be stable due to nearby jets → remove all objects from list, which are assigned to stable jets, and repeat algorithm
- Only slightly higher execution time compared to Midpoint Cone
- Publicly available: supported experiment unspecific code (good for comparison)
  - web page at HepForge: **http://projects.hepforge.org/siscone/**
- Collinear and infrared safe

Inclusive k_T and fast-k_T



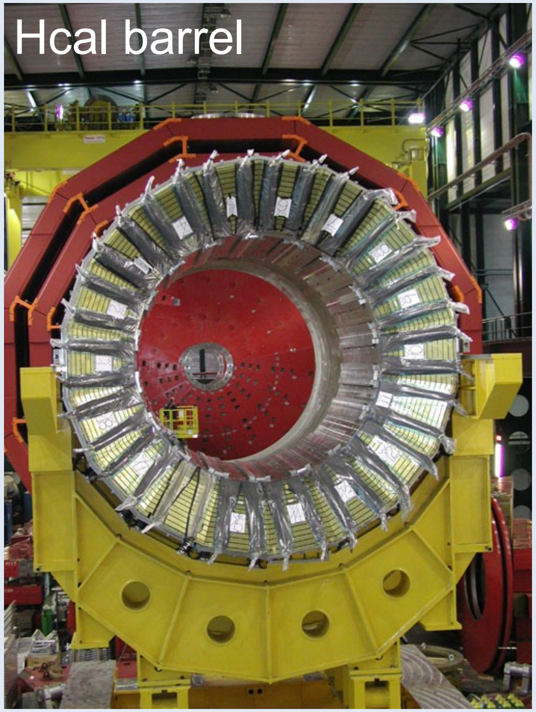
- Sequential clustering algorithm (no fixed cone size)
- For each input object calculate the "distance to the beam line"

$$d_i = \left(E_{T,i}\right)^2 \cdot D^2$$

- and the "distance" to the other particles $d_{ij} = \min\left\{E_{T,i}^2, E_{T,j}^2\right\} \cdot R_{ij}^2$
- Find for each i smallest d_{ij}
- If $d_i > d_{ij}$ move object i to the list of final jets, else merge i and j
- Advantage:
  - infrared and collinear safe
  - no fixed cone → better clustering of heavy highly boosted decaying particles
- But: High computational cost $O(N^3)$ ... ~ 10 sec for 2000 calorimeter cells
- **fast k_T:** it's enough to calculate d_{ij} to nearest neighbor: $O(N \log N)$

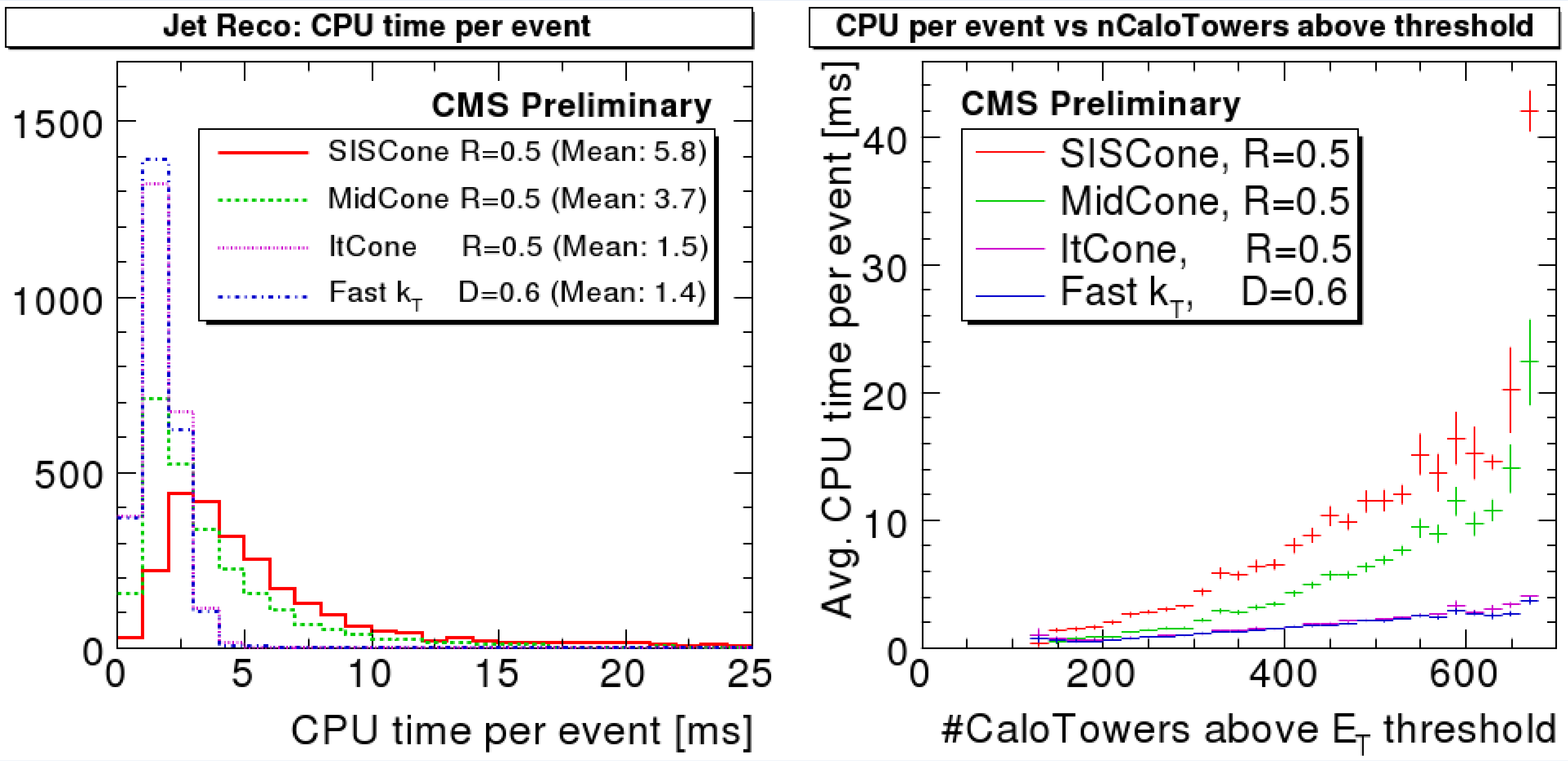
CMS Calorimeter

- ECal: Lead-Tungstate $PbWo^4$ crystals
  - Coverage: $|\eta| < 3$
  - High granularity $\Delta\eta\times\Delta\varphi = 0.0175\times0.0175$ (~80k crystals)
  - ~26 X_0
- HCal: Copper (brass) / scintillator sampling calorimeter
  - Coverage: $|\eta| < 5$
  - Granularity (barrel) $\Delta\eta\times\Delta\varphi = 0.087\times0.087$ (~4k cells)
  - ~7-11 λ (+ 4 λ from HO in barrel)
- Calorimeter towers
  - One HCal cell and 5x5 crystals
  - Overall tower threshold of 0.5 GeV
  - 82 towers in η and 72/36/18 towers in φ



Comparison: Computing Time

- Similar computing times for all algorithms (~0.002 sec) → short compared to total reconstruction time (~10 sec)
- 10% faster SIS Cone adapted in near future

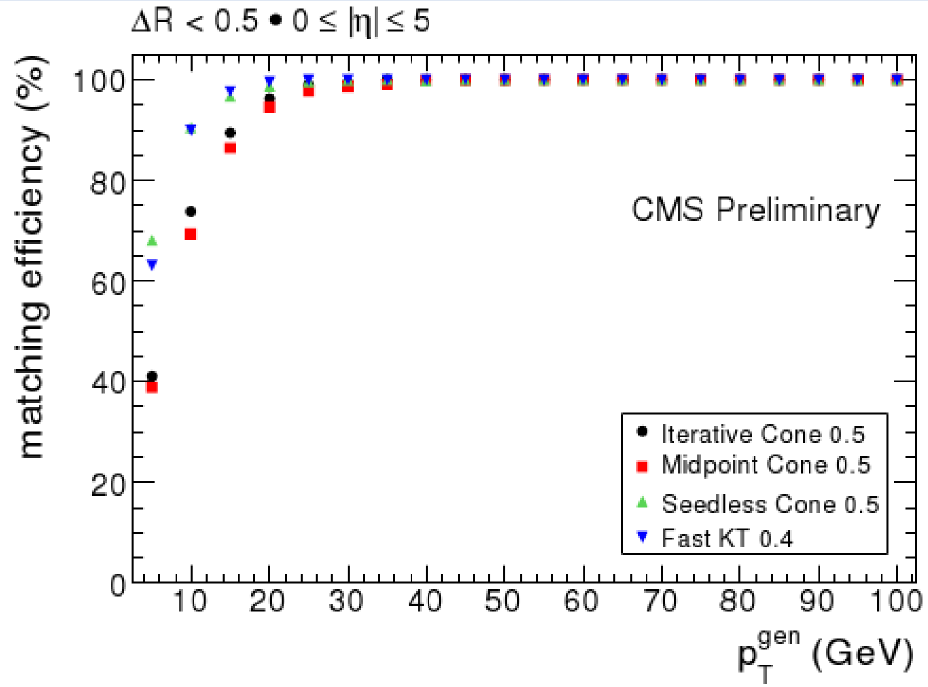


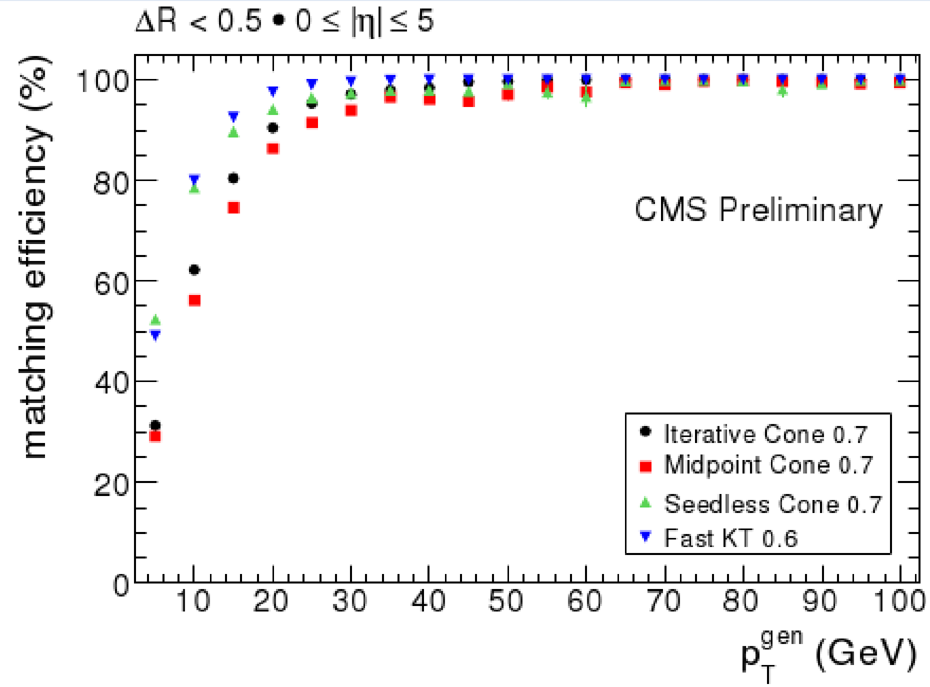
Comparison: Matching Efficiency

- One meaningful measure of reconstruction efficiency

$$\epsilon = \frac{\text{particle jets matched to Calo jet within } \Delta R = 0.5}{\text{total number of particle jets}}$$

- Strong dependence on position resolution, ΔR and cone size
- SIS Cone and k_T tend to have better performance than Midpoint and Iterative Cone
- Smaller cone size parameter → better matching efficiency

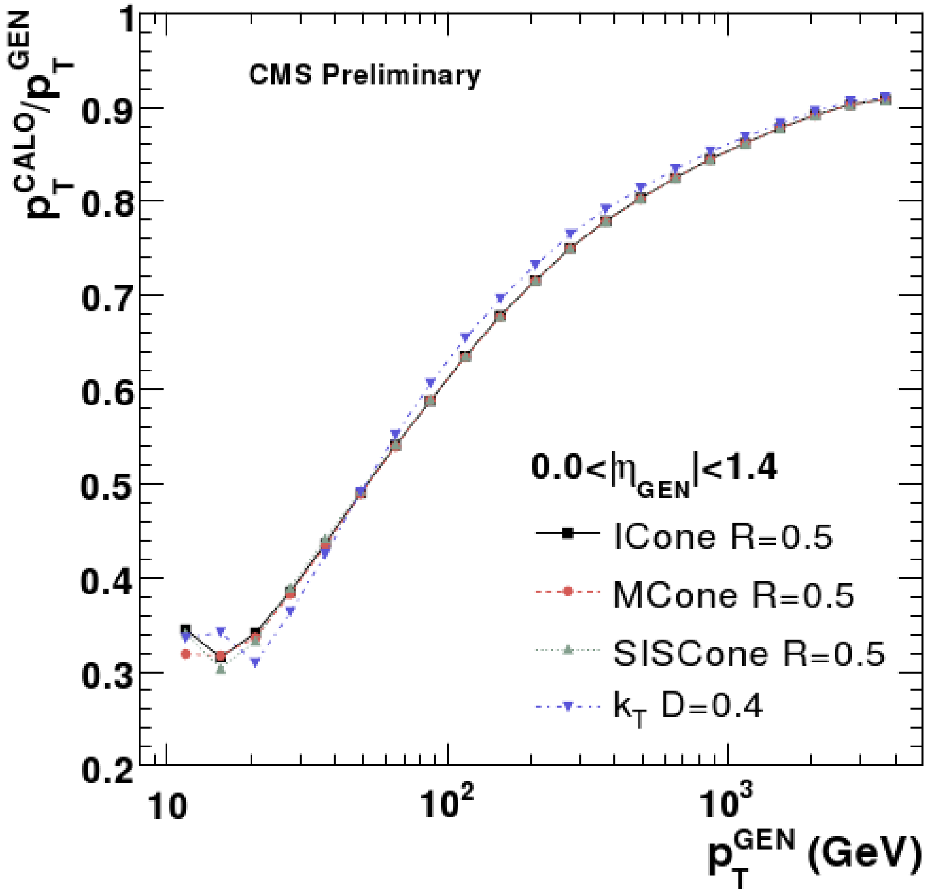


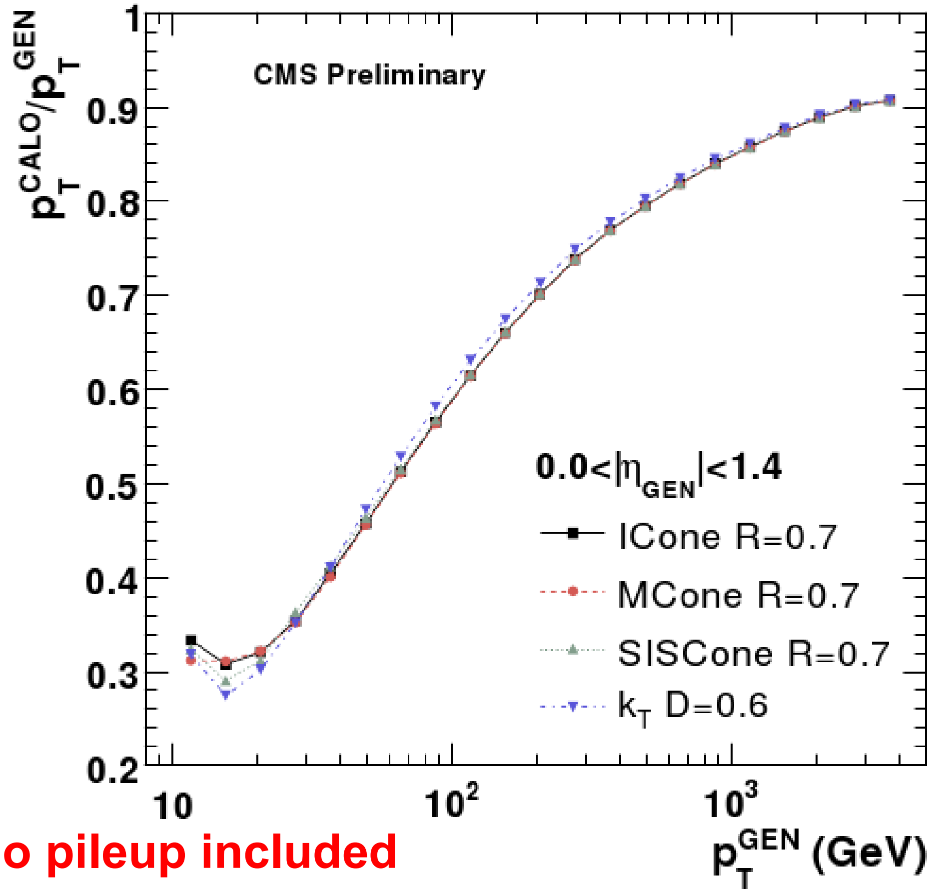


Comparison: Response

$$\text{Response} = \frac{\text{Calorimeter jet } p_T}{\text{Particle jet } p_T}$$

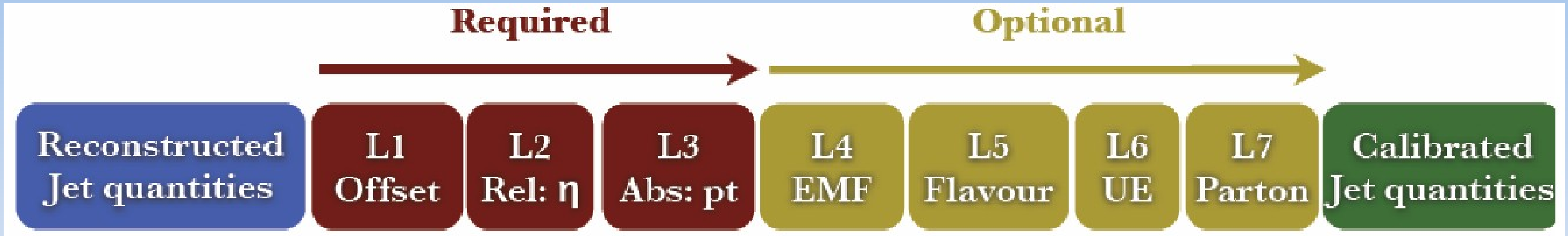
- Good agreement of R=0.5 / D=0.4 as well as R=0.7 / D=0.6 for all detector regions
- MC truth based corrections; one single calibration step



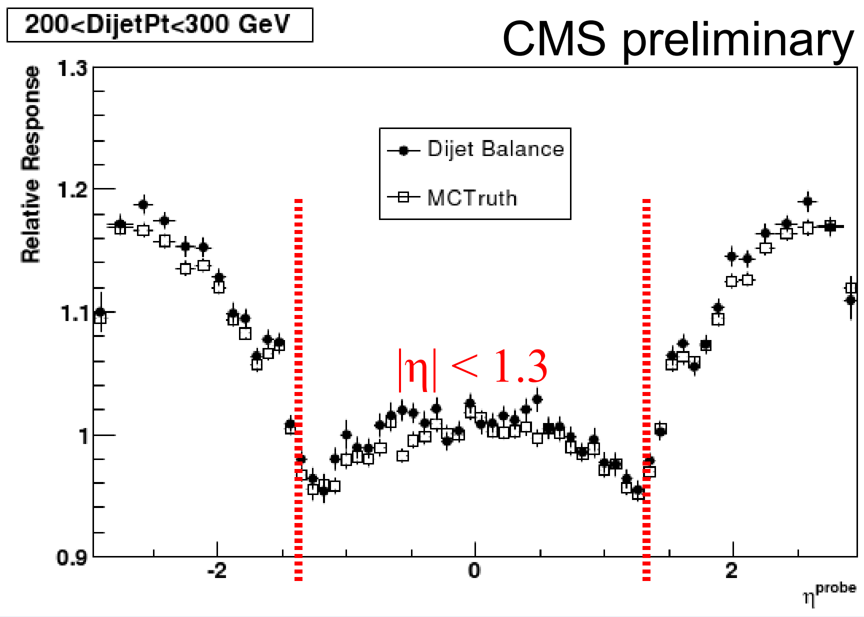


no pileup included

Jet Calibration Plans



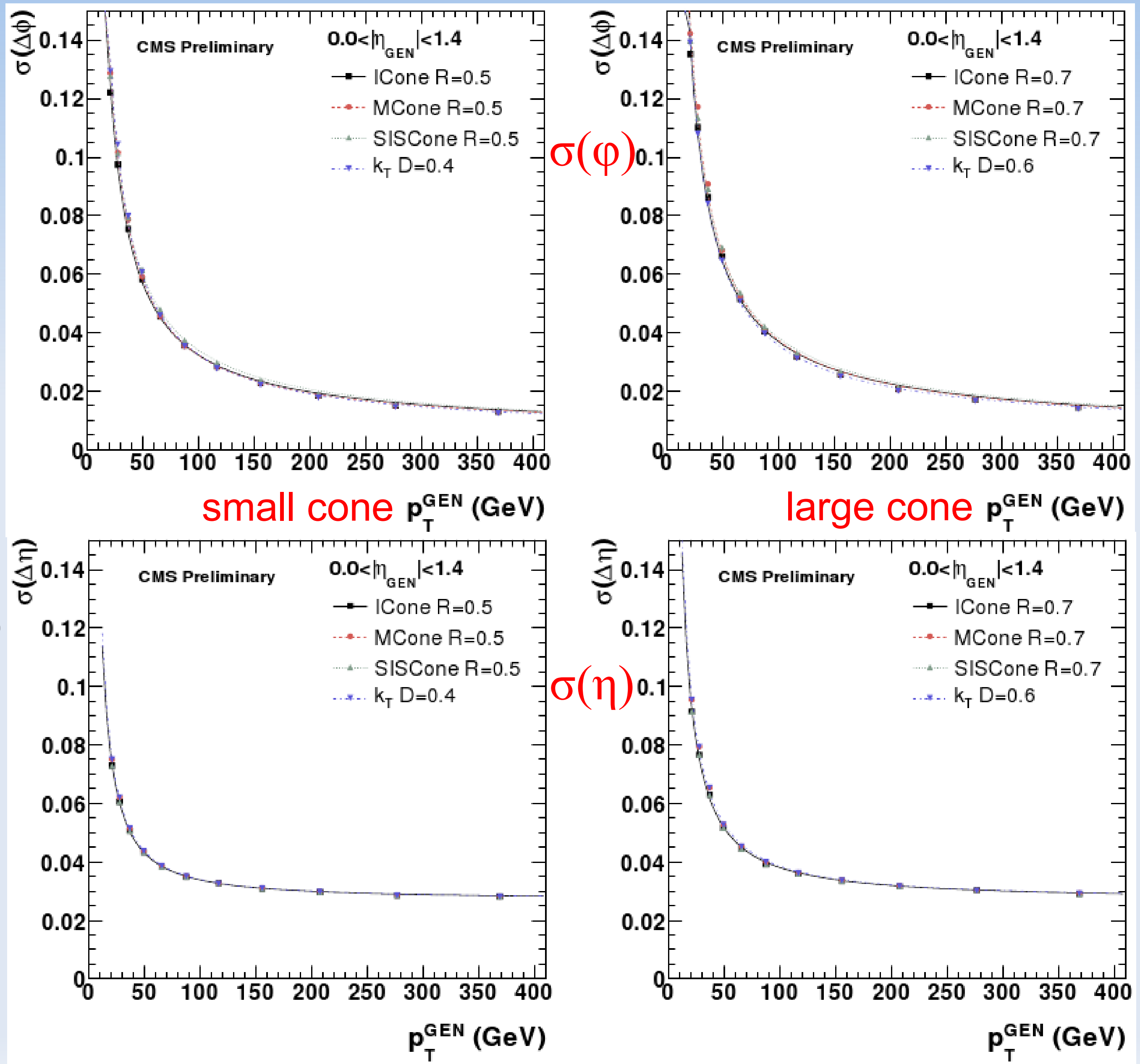
- Corrections factorized by residual corrections
- Factorization facilitates the use of data driven corrections; breaking the correction into pieces that are naturally measured in colliders
  - L1: pile-up and noise measured in zero-bias events
  - L2: jet response vs. η relative to barrel using dijet balance etc.
  - L3: jet response vs. p_T found in barrel using γ/Z + jets, top etc.
- Allows data-driven corrections as they emerge to easily replace MC truth



jet response relative to η for $|\eta|<1.3$
Good agreement for data driven dijet-balance and MC truth (closure test)

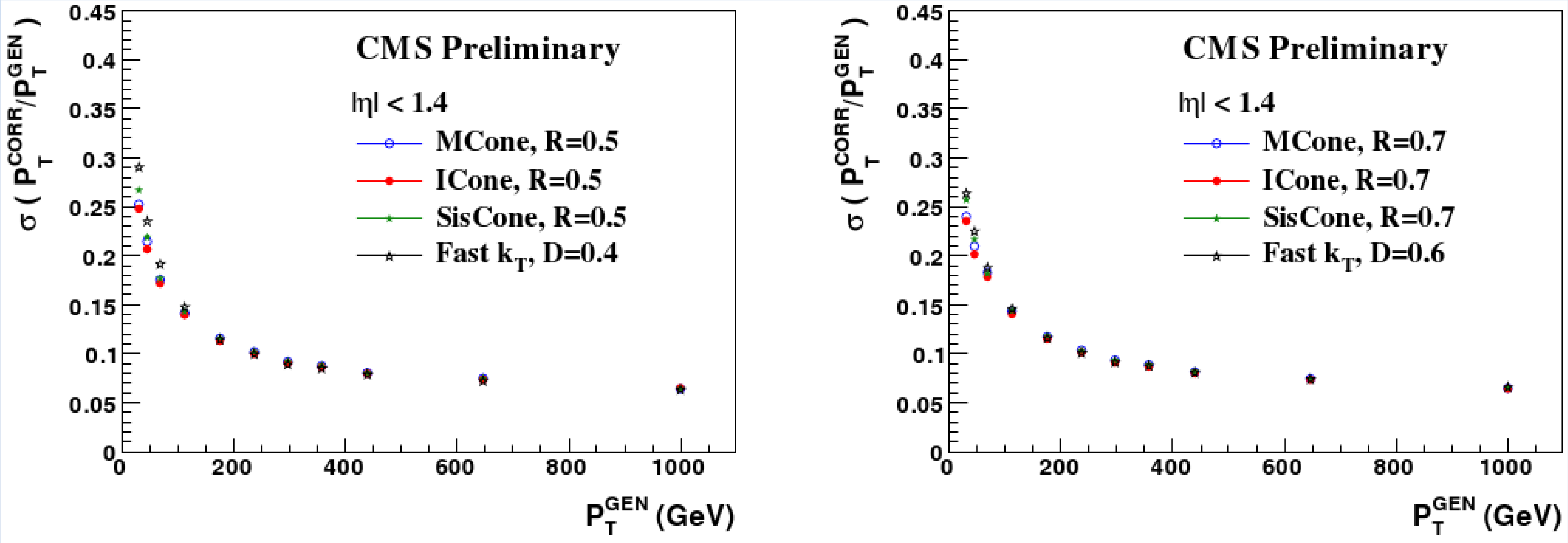
Comparison: η and φ Resolution

- η and φ Resolution almost independent on jet algorithm and cone sizes in barrel region (slightly better angular resolution for small cone size parameters)
- Dependence might be different in forward region (different granularity than in barrel)
- Shown results assume vertex position at $z = 0$ $\rightarrow$ resolution are slightly overestimated



Comparison: Energy Resolution

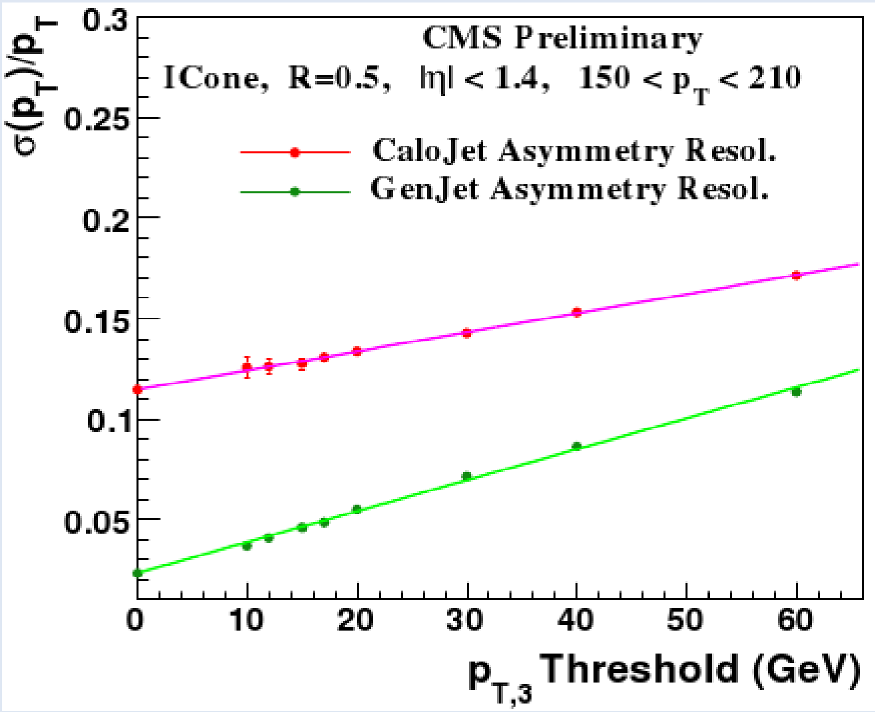
- Energy resolution of MC corrected jets almost independent on cone size parameters
- k_T algorithm slightly worse at small p_T

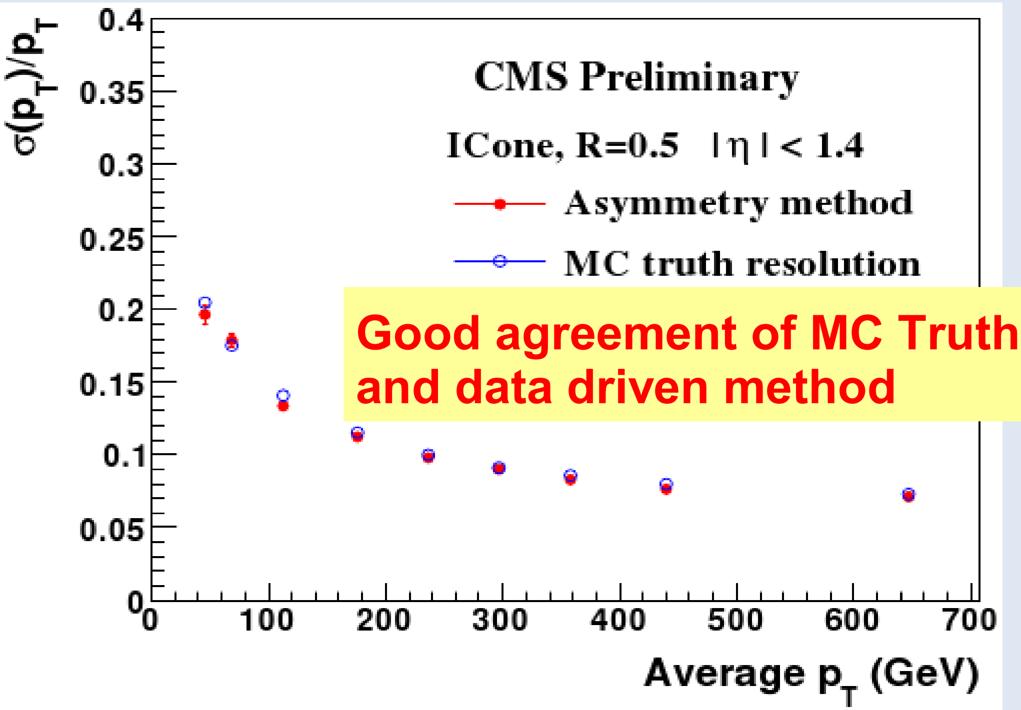


- Only calorimeter jets matched to gen jets within $R = 0.3$ are used
- Resolution determined by fit within 2.5 RMS

Resolution Measurement from Data

- Asymmetry method with dijet events $$A = \frac{p_T^{\text{jet1}} - p_T^{\text{jet2}}}{p_T^{\text{jet1}} + p_T^{\text{jet2}}}$$
- Jet resolution σ_{jet} related to σ_A $$\frac{\sigma_{p_T}}{p_T} = \sqrt{2}\,\sigma_A$$
- Dijet events often have additional jet from soft radiation
- Study resolution as function of cut on third jet → extrapolate cut to $p_T = 0$ GeV

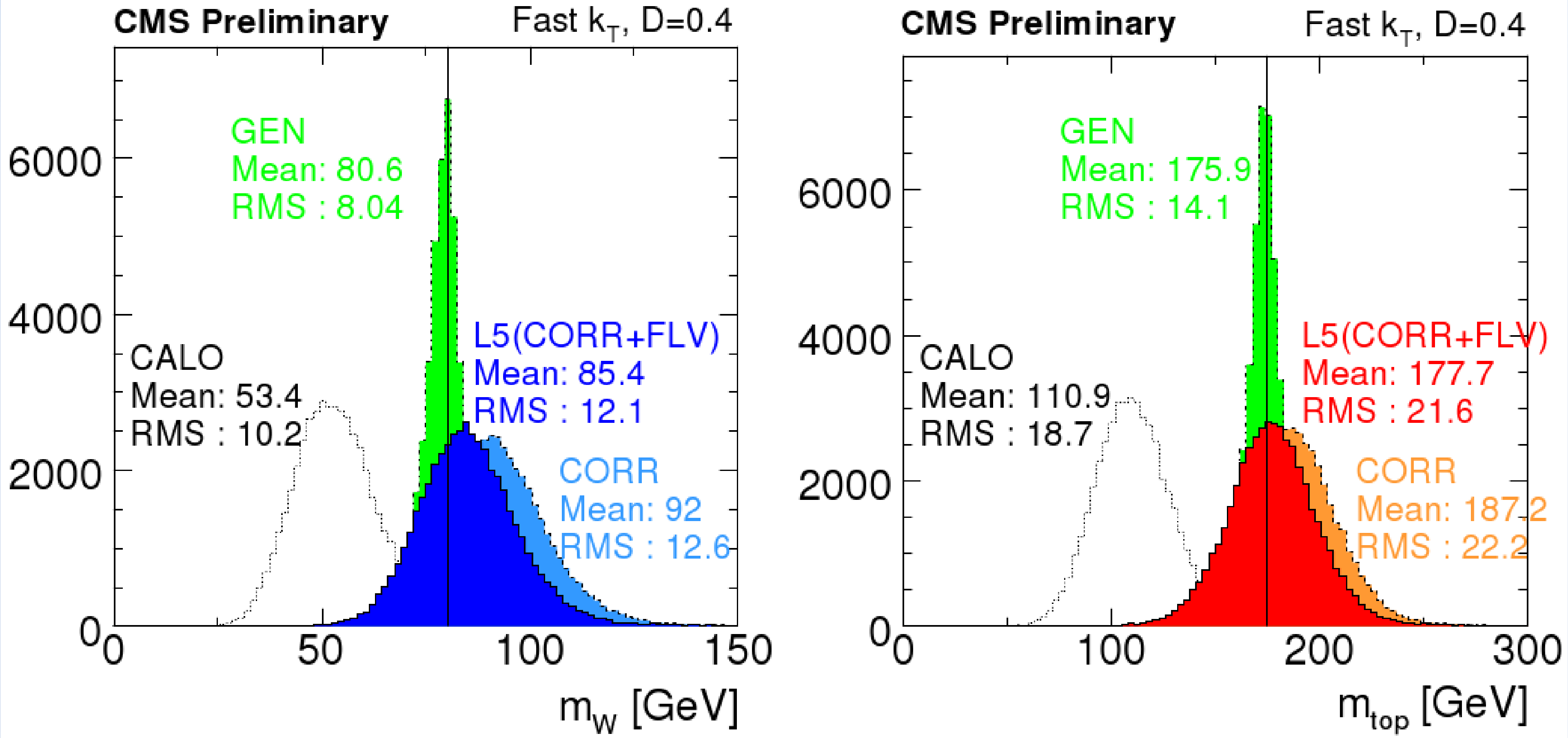




Good agreement of MC Truth and data driven method

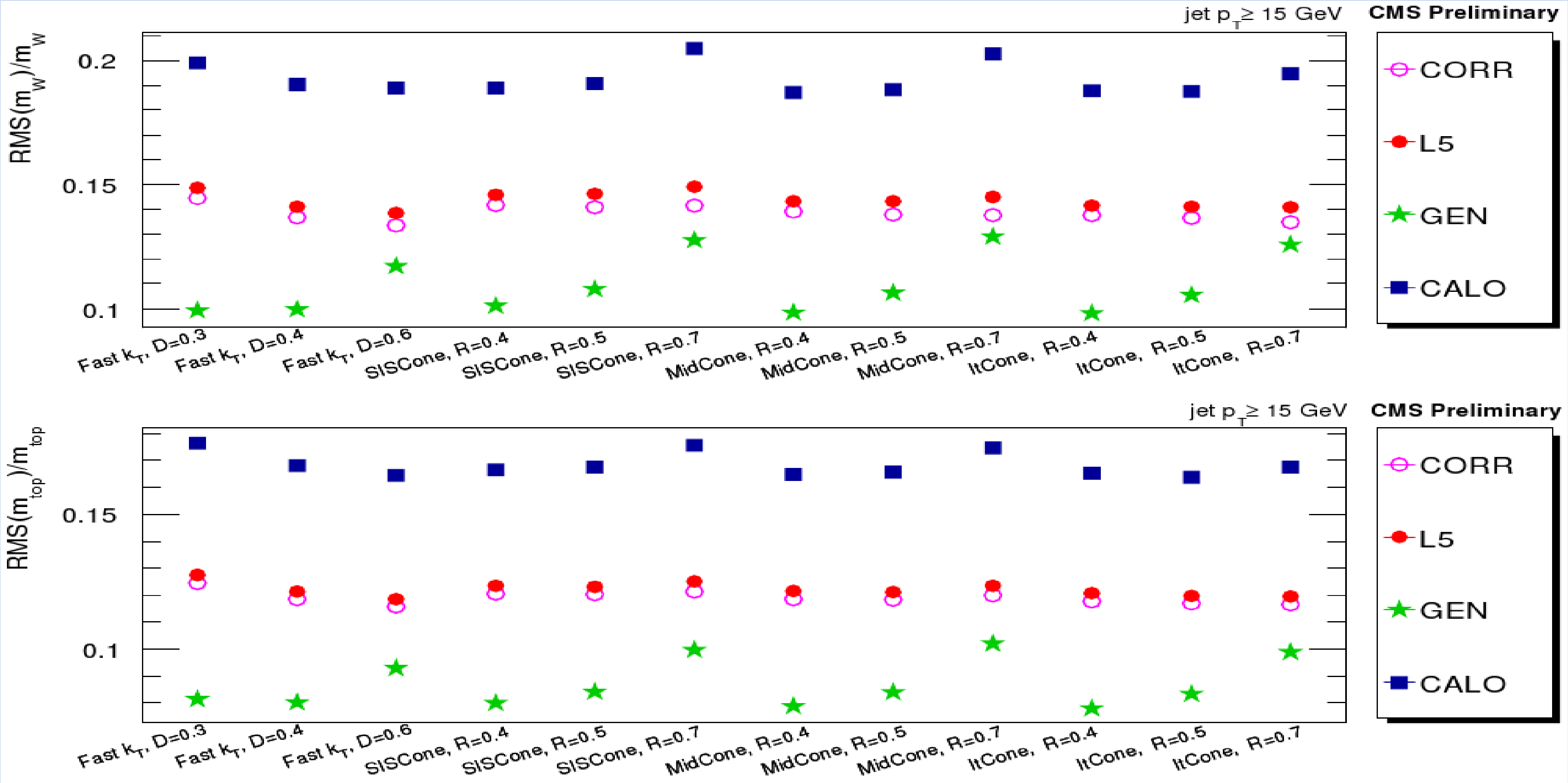
Performance in ttbar-Events

- Use independent $t\bar{t}$ events to validate jet performance
- MC corrections with additional flavour dependent corrections (L5) improve peak position of m_t and m_W significantly



Performance in ttbar-Events

Overall good agreement in mass resolution for different jet algorithms → **No need to use theoretical not preferred algorithm like Iterative Cone**



Summary

- CMS supports Iterative Cone, SIS Cone and k_T
- Midpoint Cone is available as well but not reconstructed by default
- Performance of different algorithms were studied: matching efficiency, energy and position resolution; data driven methods are developed to measure this quantities
- SIS Cone has good performance, needs reasonable computing time and is collinear and infrared safe; overall good agreement between different algorithms → CMS proposes SIS Cone as standard cone based algorithm
- MC jet corrections available and data driven methods are on the way
- Additional techniques like **particle flow**, **jet + tracks** or **track only jets** are also studied at CMS → benefit from accurate tracker information

Thanks to the CMS JetMET group

Reference: *Performance of Jet Algorithms in CMS* (CMS PAS JME-07-003)

Back-up slides

SIS Cone

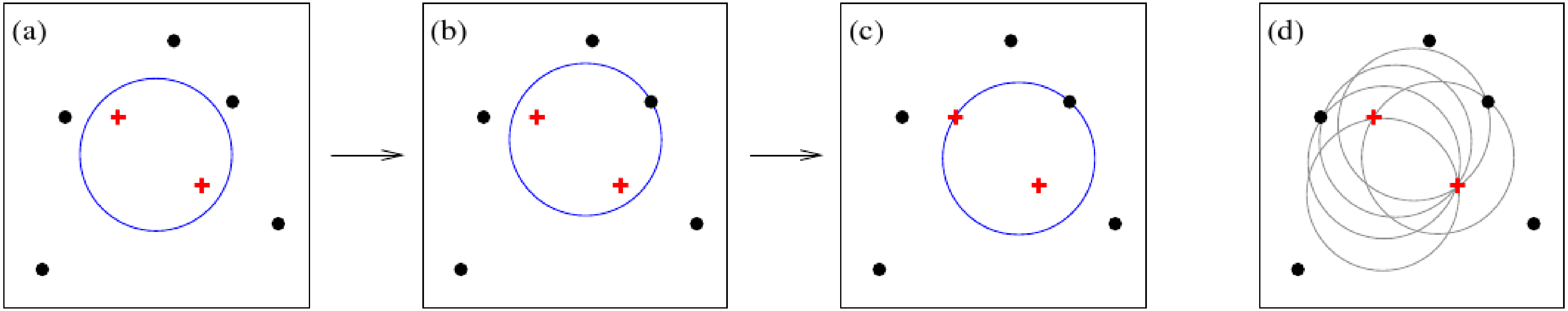
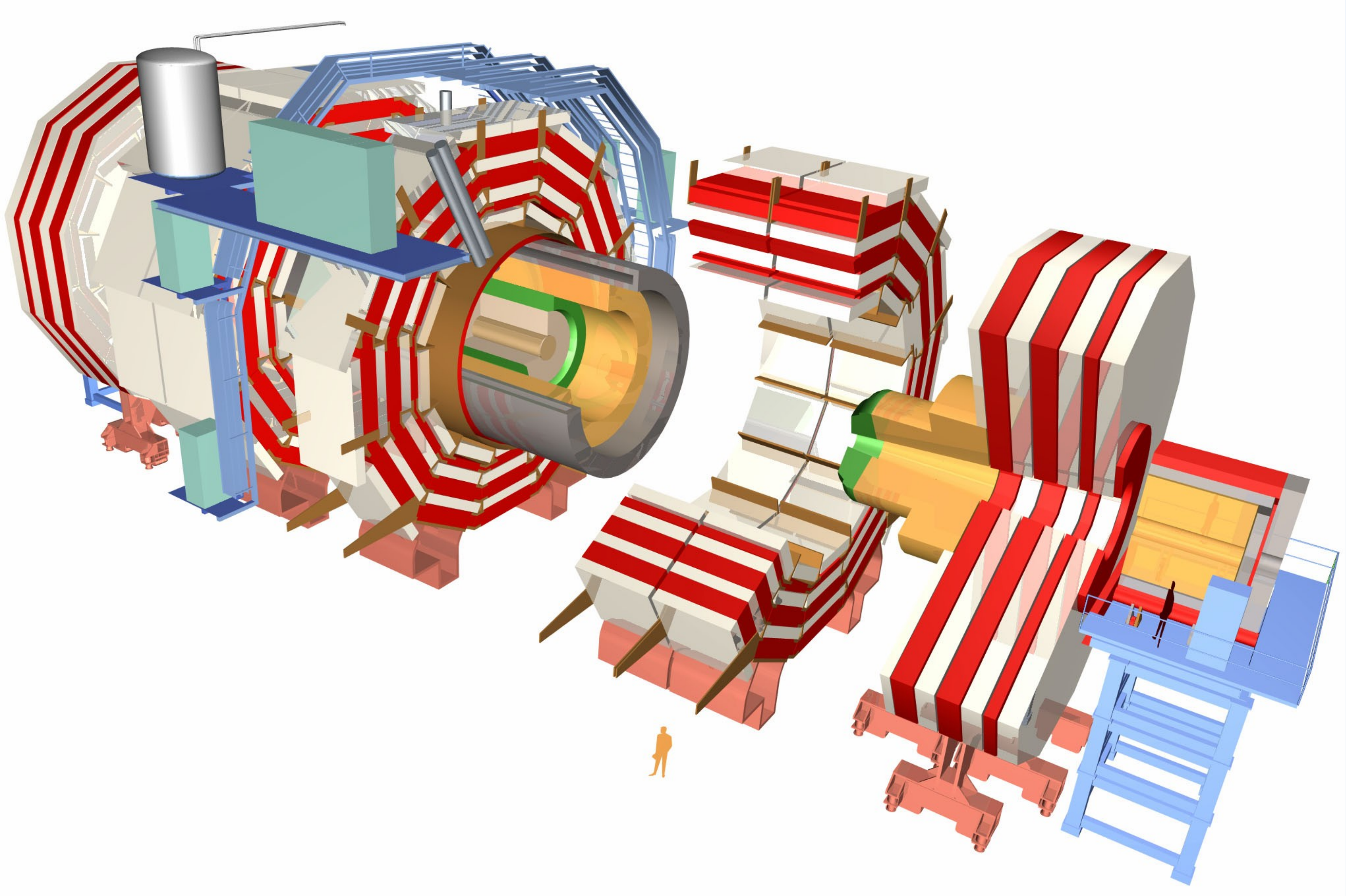


Figure 3: (a) Some initial circular enclosure; (b) moving the circle in a random direction until some enclosed or external point touches the edge of the circle; (c) pivoting the circle around the edge point until a second point touches the edge; (d) all circles defined by pairs of edge points leading to the same circular enclosure.

From G. Salam and G. Soyez hep-ph/07040292v2

The CMS Detector



CMS HCal

- Granularity decreases with increasing $|\eta|$
- Coverage up to $|\eta| \approx 5$
- Most parts inside of magnet except OuterBarell (OB)

from CMS Physics TDR Vol. II

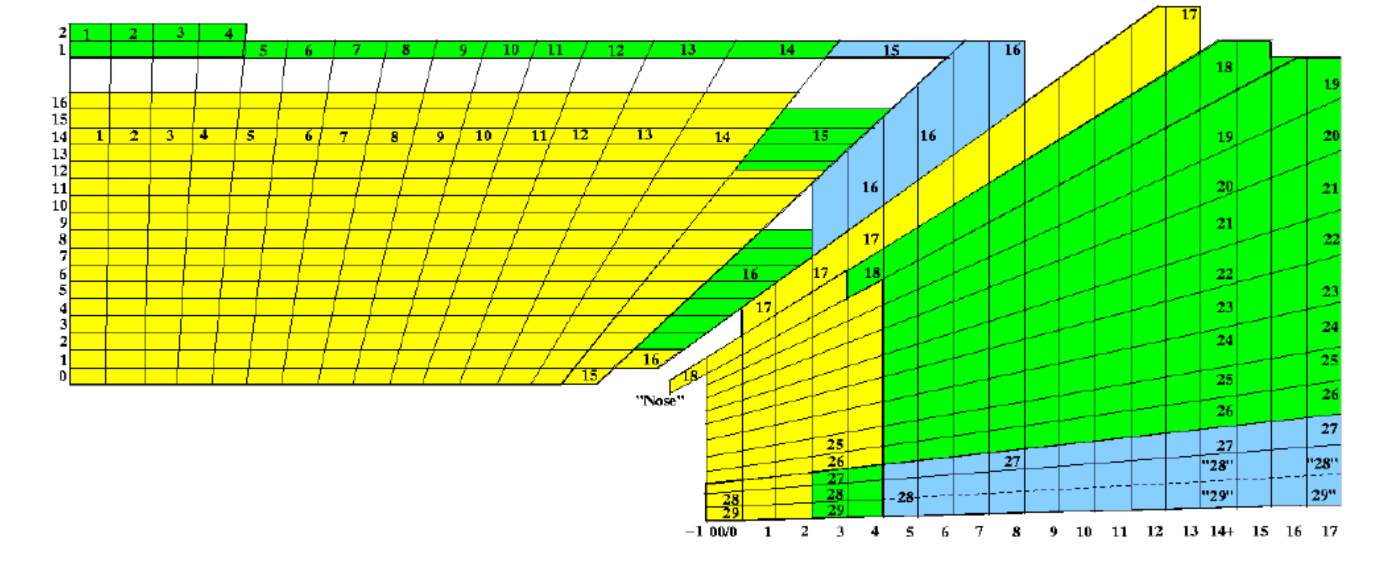


Figure 5.1: A schematic view of the tower mapping in r-z of the HCAL barrel and endcap regions.

Tower index	η range Low	η range High	Detector	Size η	Size ϕ	Depth segments
1	0.000	0.087	HB, HO	0.087	5°	HB=1, HO=1
2	0.087	0.174	HB, HO	0.087	5°	HB=1, HO=1
3	0.174	0.261	HB, HO	0.087	5°	HB=1, HO=1
4	0.261	0.348	HB, HO	0.087	5°	HB=1, HO=1
5	0.348	0.435	HB, HO	0.087	5°	HB=1, HO=1
6	0.435	0.522	HB, HO	0.087	5°	HB=1, HO=1
7	0.522	0.609	HB, HO	0.087	5°	HB=1, HO=1
8	0.609	0.696	HB, HO	0.087	5°	HB=1, HO=1
9	0.696	0.783	HB, HO	0.087	5°	HB=1, HO=1
10	0.783	0.870	HB, HO	0.087	5°	HB=1, HO=1
11	0.879	0.957	HB, HO	0.087	5°	HB=1, HO=1
12	0.957	1.044	HB, HO	0.087	5°	HB=1, HO=1
13	1.044	1.131	HB, HO	0.087	5°	HB=1, HO=1
14	1.131	1.218	HB, HO	0.087	5°	HB=1, HO=1
15	1.218	1.305	HB, HO	0.087	5°	HB=2, HO=1
16	1.305	1.392	HB, HE	0.087	5°	HB=2, HE=1
17	1.392	1.479	HE	0.087	5°	HE=1
18	1.479	1.566	HE	0.087	5°	HE=2
19	1.566	1.653	HE	0.087	5°	HE=2
20	1.653	1.740	HE	0.087	5°	HE=2
21	1.740	1.830	HE	0.090	10°	HE=2
22	1.830	1.930	HE	0.100	10°	HE=2
23	1.930	2.043	HE	0.113	10°	HE=2
24	2.043	2.172	HE	0.129	10°	HE=2
25	2.172	2.322	HE	0.150	10°	HE=2
26	2.322	2.500	HE	0.178	10°	HE=2
27	2.500	2.650	HE	0.150	10°	HE=3
*28	2.650	3.000	HE	0.350	10°	HE=3
29	2.853	2.964	HF	0.111	10°	HF=2
30	2.964	3.139	HF	0.175	10°	HF=2
31	3.139	3.314	HF	0.175	10°	HF=2
32	3.314	3.489	HF	0.175	10°	HF=2
33	3.489	3.664	HF	0.175	10°	HF=2
34	3.664	3.839	HF	0.175	10°	HF=2
35	3.839	4.013	HF	0.174	10°	HF=2
36	4.013	4.191	HF	0.178	10°	HF=2
37	4.191	4.363	HF	0.172	10°	HF=2
38	4.363	4.538	HF	0.175	10°	HF=2
39	4.538	4.716	HF	0.178	10°	HF=2
40	4.716	4.889	HF	0.173	20°	HF=2
41	4.889	5.191	HF	0.302	20°	HF=2

Jet Calibration – MC correction

- Monte Carlo corrections determined by the non-linear equation

$$\text{Calo jet } E_T(\eta) = \text{particle jet } E_T \times \text{Response}(\text{Calo jet } E_T, \eta)$$

- Iterative solution starting with

$$k(\text{Calo jet } E_T, \eta) = \frac{1}{\text{Response}(\text{Calo jet } E_T, \eta)}$$

- Calorimeter jet 4-vectors p have to be multiplied with k

$$p' = k \cdot p$$

- Response is a function of particle jet E_T → iterating till corrected Calorimeter jet E_T converges to particle Jet E_T
- i^{th} iteration ($k_0 = 1$): $$k(\text{Calo jet } E_T, \eta) = \frac{1}{\text{Response}(\text{Calo jet } E_T \times k_{i-1}, \eta)}$$
- **General remarks:**
  - Larger jet size parameter tend to requires larger corrections
  - Differences among comparable algorithms are pronounced at low E_T
  - As $|\eta|$ increases differences among jets with different sizes become larger