Pileup subtraction at the particle level

Peter Berta[1], **Martin Spousta**[1], David Miller[2]

[1] Charles University in Prague
[2] Enrico Fermi Institute, University of Chicago

20th Particles and Nuclei International Conference, 25-29 August 2014, Hamburg, Germany

Outline

- Introducing the problem of pileup
- Brief overview of existing methods and a summary of latest developments
- Constituent subtraction
- Performance of the constituent subtraction on jets and jet shapes
- Full Event subtraction
- Conclusions

Introduction

- Pileup (=superposition of multiple pp collisions at each bunch spacing) in **pp collisions**. It affects:
  - jets and jet shapes variables, e.g. used to tag boosted objects
  - missing-pt and event shapes
  - photons and lepton isolation
- Underlying event in **heavy ion collisions**. It affects:
  - jets and jet internal structure used to study jet quenching
  - photons and lepton isolation
- Backgrounds at LHC influence practically each measurement – starting from Higgs ending with jet quenching => need to develop (or keep improving) methods for subtraction!

mean number of collisions

energy underneath the R=0.4 jet

	Pileup (μ)	Additional energy
LHC run II	~50	15 GeV
HL LHC	~140	40 GeV
Heavy Ions (central)	~0	>100 GeV

Basic overview of methods

Method	Works for any experiment?	Can do full event subtraction?	Can work without tracking?	Can work in heavy ions?	Is the code public?
4-mometum subtraction	✔	✘	✔	✔	✔
Shape expansion	✔	✘	✔	✘	✔
Grooming *	✔	✘	✔	✘	✔
Charged hadrons	✘ (CMS)	✔	✘	✘	-
Topoclusters	✘ (ATLAS)	✔	✔	✘	-
Cleansing **	✔	✘	✘	✘	✔
SoftKiller	✔	✔	✔	?	✔
PUPPI	✔	✔	✘	?	✘
Constituent Subtraction	✔	✔	✔	?	✔

main focus of this talk

per-particle methods

* Grooming is a general term including various methods (see backup) for details see e.g. arXiv:1311.2708.

** Cleansing is not discussed in this presentation, method is described in arXiv:1309.4777 and further discussed in arXiv:1404.7353.

Basic four-momentum subtraction

Cacciari, Salam, PLB 659 (2008) 119

- Calculate the **background density** ρ as a median over patches of transverse momentum p_T in the area A

$$\rho = \underset{i \in \text{patches}}{\text{median}} \left\{ \frac{p_{ti}}{A_i} \right\}$$

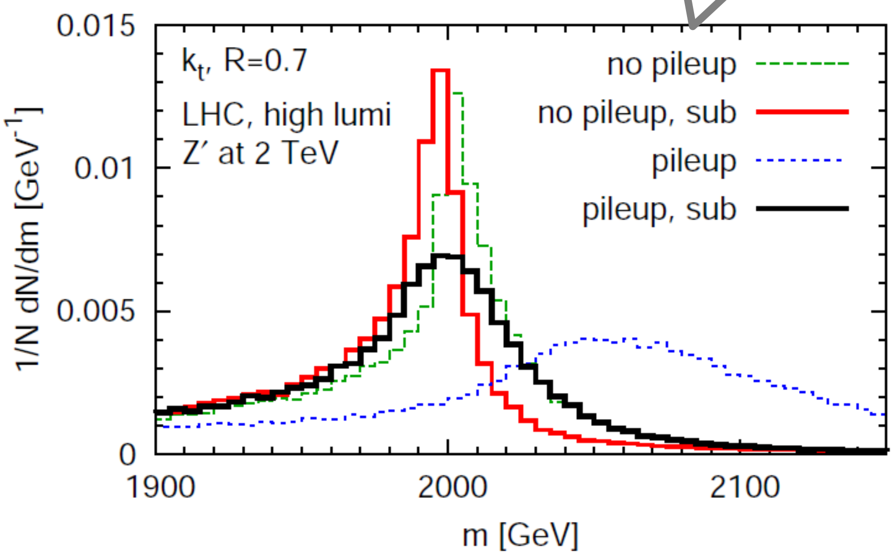
(The use of median excludes a bias from jets)

- Subtract the jet **four-momentum** as follows:

$$p_\mu^{(\text{sub})} = p_\mu - A_\mu \rho$$

- … the basic scheme for subtraction of background both in pp and HI
- It corrects the kinematics, but it **cannot correct the internal structure** of the jet

Demonstration of performance on Z' mass reconstructed from di-jets

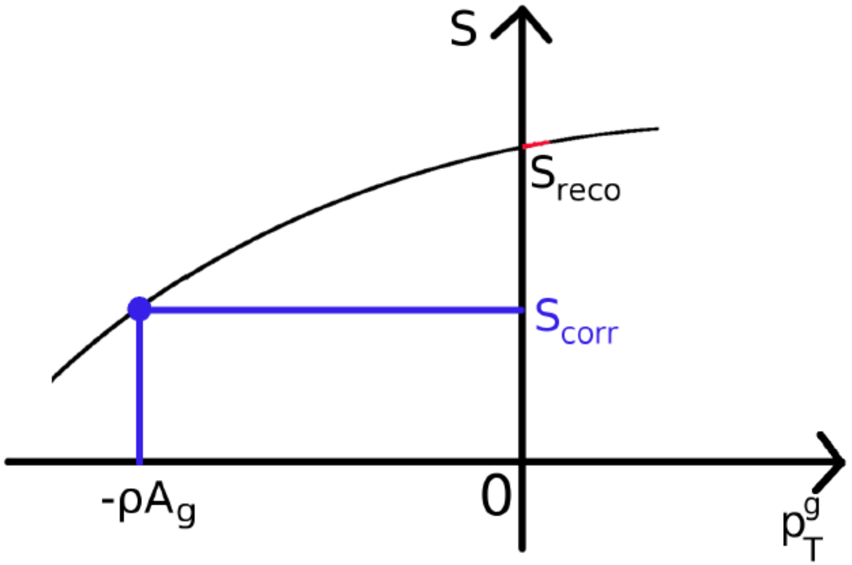


Shape expansion method

Soyez, et al. PRL 110 (2013) 162001

- Basic objects: **ghosts** = infinitesimally soft mass-less particles uniformly covering the η, ϕ space. Each ghost covers certain area A_g.
- Ghosts are clustered by the jet algorithm with real particles with no impact on kinematics of a jet.
- To subtract background, one might set for each ghost $p_{T,g} = -\rho A_g$
- It is cumbersome to work with negative particles => do **extra-polation** using Taylor expansion.
- Jet shape *S* = arbitrary function of jet constituents. Corrected jet shape:

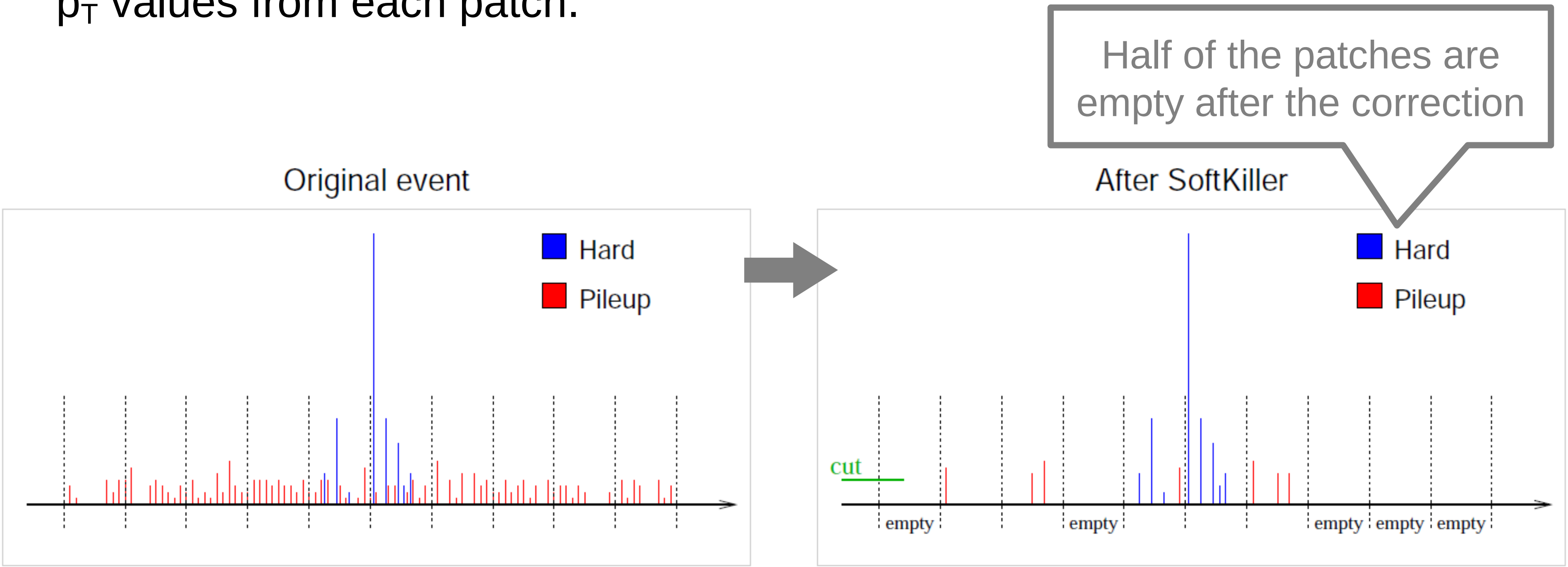
$$S_{\mathrm{corr}} = \sum_{k=0}^{\infty} (-\rho A_g)^k \cdot \frac{\partial^k S(p_{\mathrm{T}}^g)}{\partial {p_{\mathrm{T}}^g}^k}\bigg|_{p_{\mathrm{T}}^g=0}$$



SoftKiller

Cacciari et al. arXiv:1407.0408

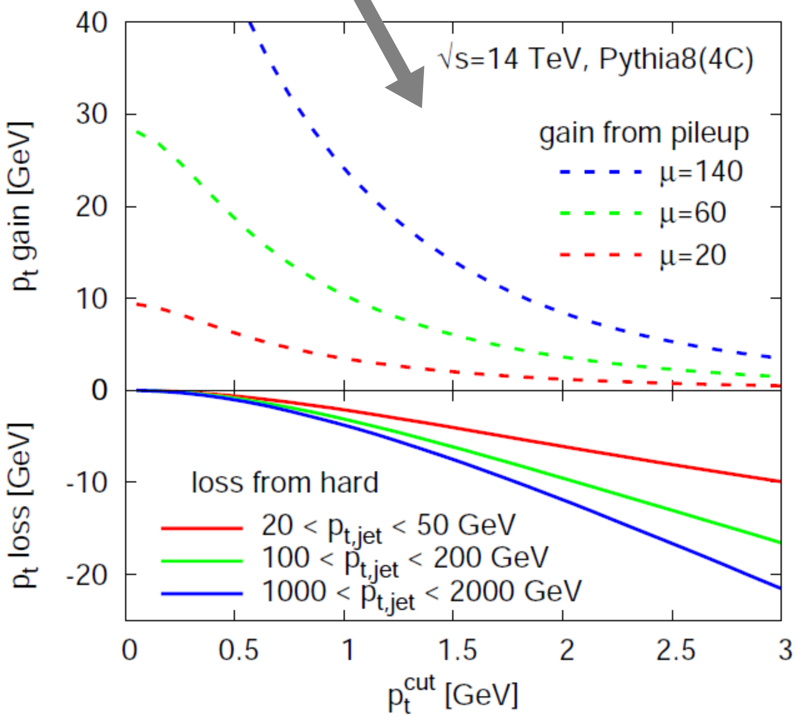
- Discards particles that fall below a certain transverse momentum threshold.
- Estimate the threshold event-by-event such that ρ estimated from the median-area method is evaluated to be zero.
- Technically very simple – just need to find a median from set of maximal p_T values from each patch.



SoftKiller

Cacciari et al. arXiv:1407.0408

- Biases on jet energy scale: net positive bias from background particles that are not subtracted, net negative bias from particles lost from jets.
- For the typical p_T cuts and grid spacing (=size of patches) of a~0.4 the biases cancel.



+ very fast and simple => e.g. potential for triggering

? may induce a flavor dependence of jet energy scale

? performance for larger jet radii and jet shapes

PUPPI

(PileUp Per Particle Identification)

Bertolini et al., arXiv:1407.6013

- For each particle define a local shape variable α_i using other particles

$$\alpha_i = \log \sum_{j \in \text{event}} \xi_{ij} \times \Theta(R_{\min} \le \Delta R_{ij} \le R_0)$$

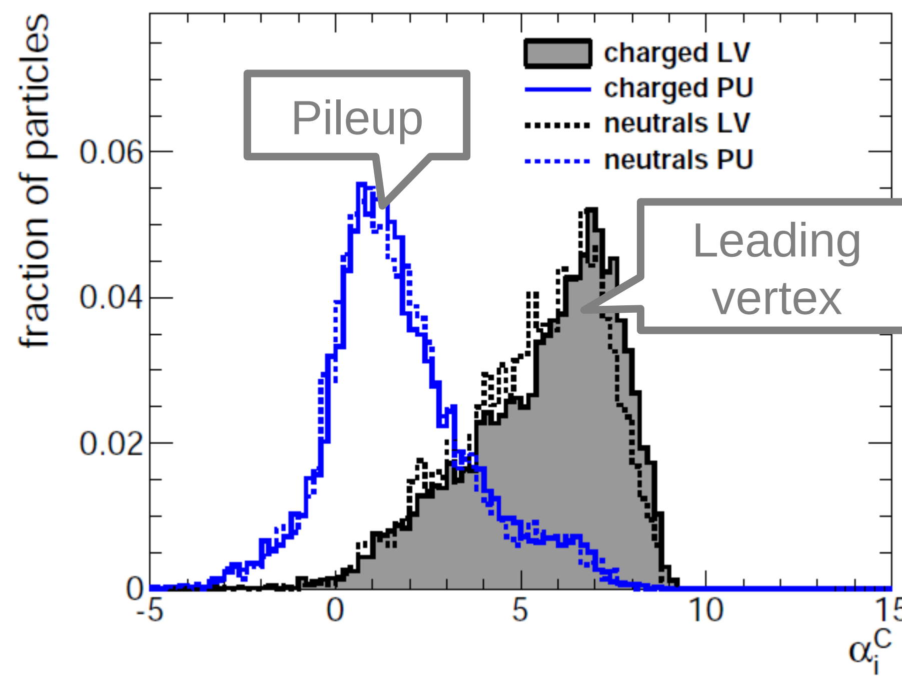
do it within a cone of certain size

$$\xi_{ij} = \frac{p_{Tj}}{\Delta R_{ij}}$$

High-pt particles are likely from jets

Parton radiation is predominantly collinear

- Then translate the local shape variable to particle weights that multiply the track momentum.



PUPPI
(PileUp Per Particle Identification)

Bertolini et al., arXiv:1407.6013

- Translate the local shape variable to particle weights using cumulative distribution function of χ^2 distribution

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

$$\chi_i^2 = \Theta(\alpha_i - \bar{\alpha}_{\mathrm{PU}}) \times \frac{(\alpha_i - \bar{\alpha}_{\mathrm{PU}})^2}{\sigma_{\mathrm{PU}}^2}$$

pileup median

=1 if a_i > median
=0 otherwise

- Weights with values between 0 and 1 used to multiply the p_T of particle.
- Zero weight discards the particle.
- Other discriminants can be easily included.

Constituent subtraction

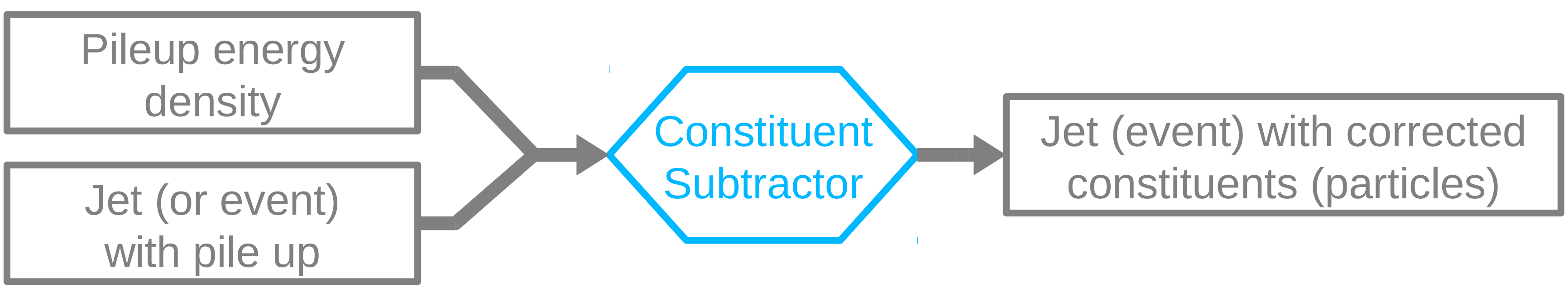
- Two main ingredients of constituent subtraction:
  - background (pileup) p_T density

$$\rho = \text{median}_{\text{patches}} \left\{ \frac{p_{\text{T,patch}}}{A_{\text{patch}}} \right\}$$

  - ghosts

$$p^g_\mu = p^g_{\text{T}} \cdot [\cos\phi^g, \sin\phi^g, \sinh y^g, \cosh y^g]$$

- Constituent subtraction provides a rule for associating the background p_T density with a given constituent, independent of any tracking information.



Constituent subtraction algorithm

- For each event:
  1. estimate the background p_T density, ρ
  2. add ghosts ($p_{T,g}$ –> 0) among particles and run the jet algorithm
- For each jet in the event:
  3. for each ghost set $p_{T,g} = \rho A_g$
  4. evaluate distance $\Delta R_{i,k}$ between particle *i* and ghost *k* for all particle-ghost pairs

$$\Delta R_{i,k} = p_{\mathrm{T}i}^{\alpha} \cdot \sqrt{\left(y_i - y_k^g\right)^2 + \left(\phi_i - \phi_k^g\right)^2}$$

  5. sort the distances and iteratively change the momenta as follows

$$\text{If } p_{\mathrm{T}i} \geq p_{\mathrm{T}k}^g: \quad p_{\mathrm{T}i} \longrightarrow p_{\mathrm{T}i} - p_{\mathrm{T}k}^g, \quad p_{\mathrm{T}k}^g \longrightarrow 0; \qquad \text{otherwise:} \quad p_{\mathrm{T}i} \longrightarrow 0, \quad p_{\mathrm{T}k}^g \longrightarrow p_{\mathrm{T}k}^g - p_{\mathrm{T}i}.$$

  until no more pairs remain or $\Delta R_{i,k} > \Delta R^{max}$

  6. discard all particles with zero transverse momentum

Properties of the algorithm

- Generic method for the background subtraction.
- Generic = independent of jet algorithm, tracking or calorimeter information => can be used across experiments and across backgrounds.
- Output:
  a) jet with corrected constituents => can estimate unbiased jet shapes
  b) subtracted event => can improve e.g. missing-p_T reconstruction
- Other properties:
  – accounts for fluctuations of the background
  – is longitudinally invariant
  – is reasonably fast
  – incorporates the subtraction of mass density for massive particles (for simplicity not discussed here)
  – can accommodate any improved energy density estimates
  – can accommodate discriminants helping to identify pileup

Properties of the algorithm

- Free parameters:
  - Ghost area A_g - the smaller the better, but also slower.
  - Parameter α - configuration with α>0 prefers to subtract lower p_T particles.
  - Maximal distance, ΔR_{max} - restricts subtraction between ghost-particle pairs with large distance.
- Default settings **do not need to be tuned** e.g. for different jet sizes or different level of pile-up.
- Tuning of free parameters may lead to modest improvements.

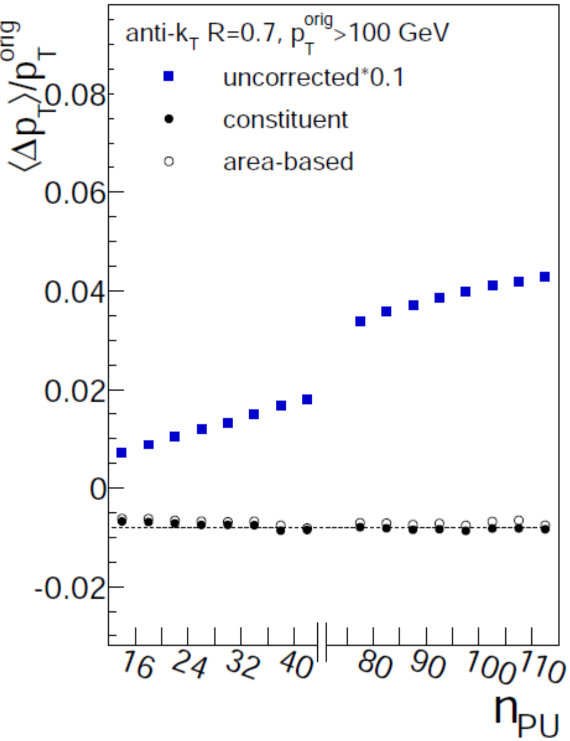
- Algorithm is publicly available from FastJet Contrib and documented in Berta et al., JHEP 06 (2014) 092.

Performance tests

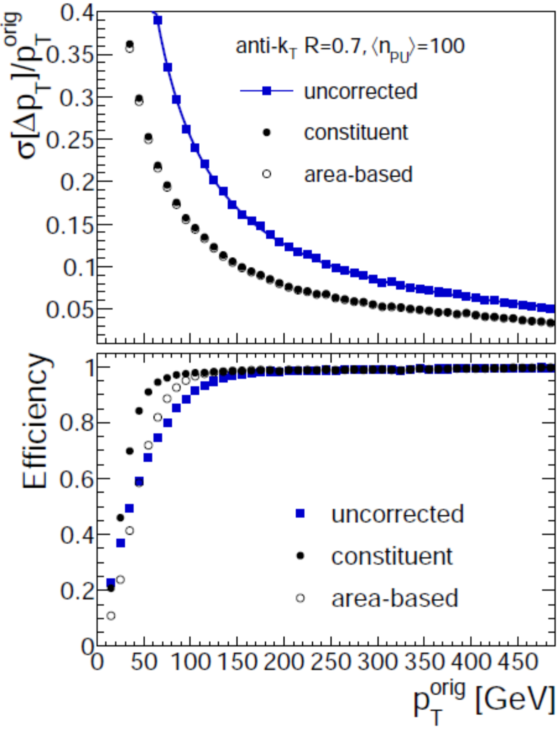
- Pythia8 simulated samples of dijets and Z' –>tt with m(Z')=1500 GeV.
- Simulation at 8 TeV without underlying event.
- Using clustering algorithms anti-kt with R=0.7, R=1.0 and Cambridge/Aachen with R = 1.2.
- Pseudorapidity cut on jets $|\eta|<2$.
- Performance compared to the area 4-vector and shape-expansion methods.

Performance – jet kinematics

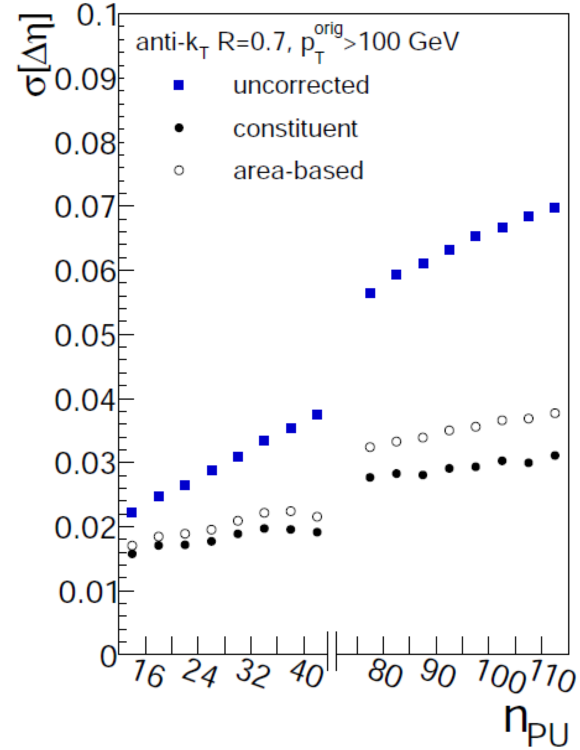
Jet energy scale



Jet energy resolution and efficiency



Jet position resolution

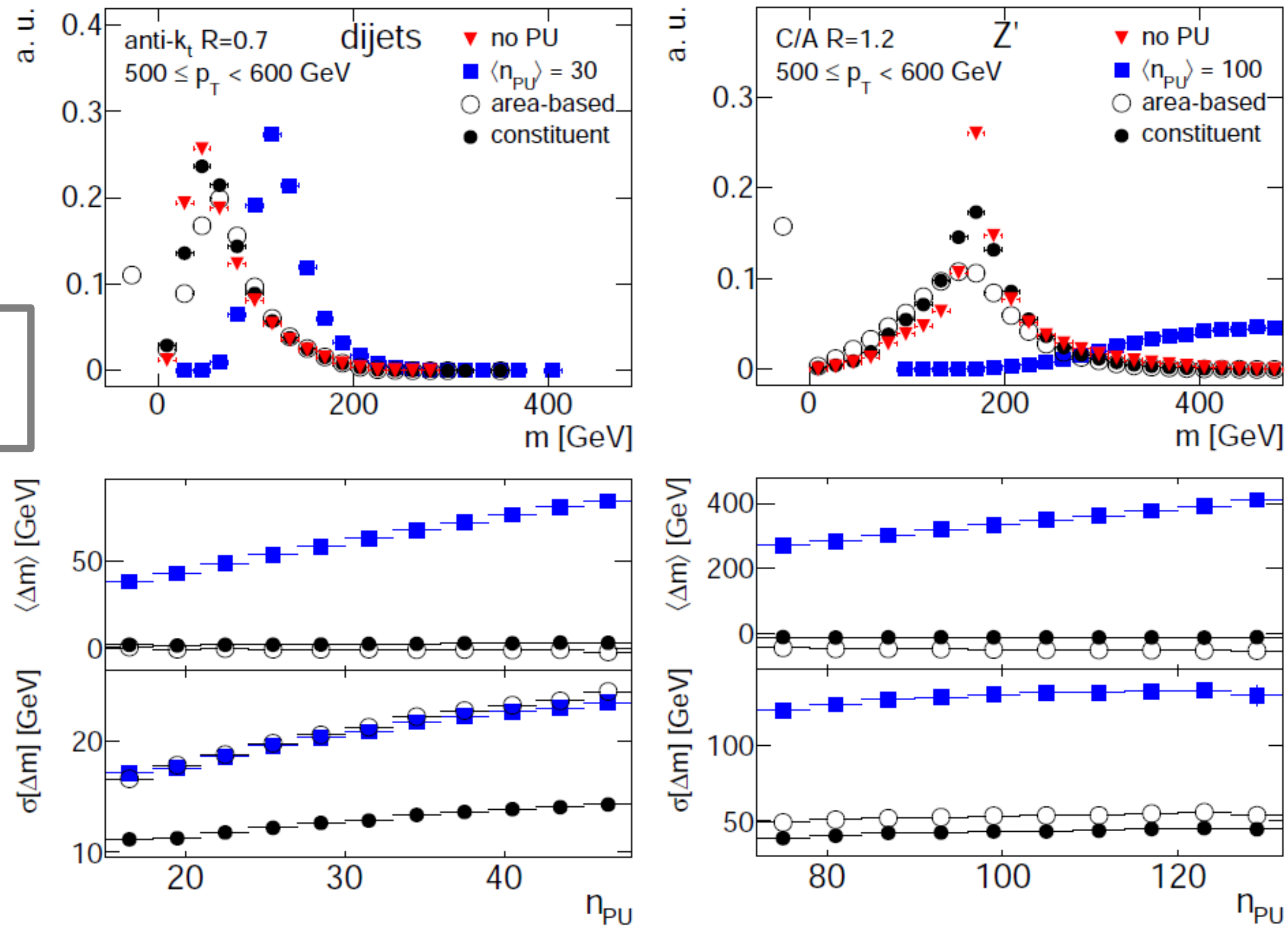


- Perfect stability of jet energy scale with increasing pileup.
- Significant improvement in jet reconstruction efficiency and jet position resolution. Jet energy resolution similar to area-based subtraction.

Performance – jet and Z' mass

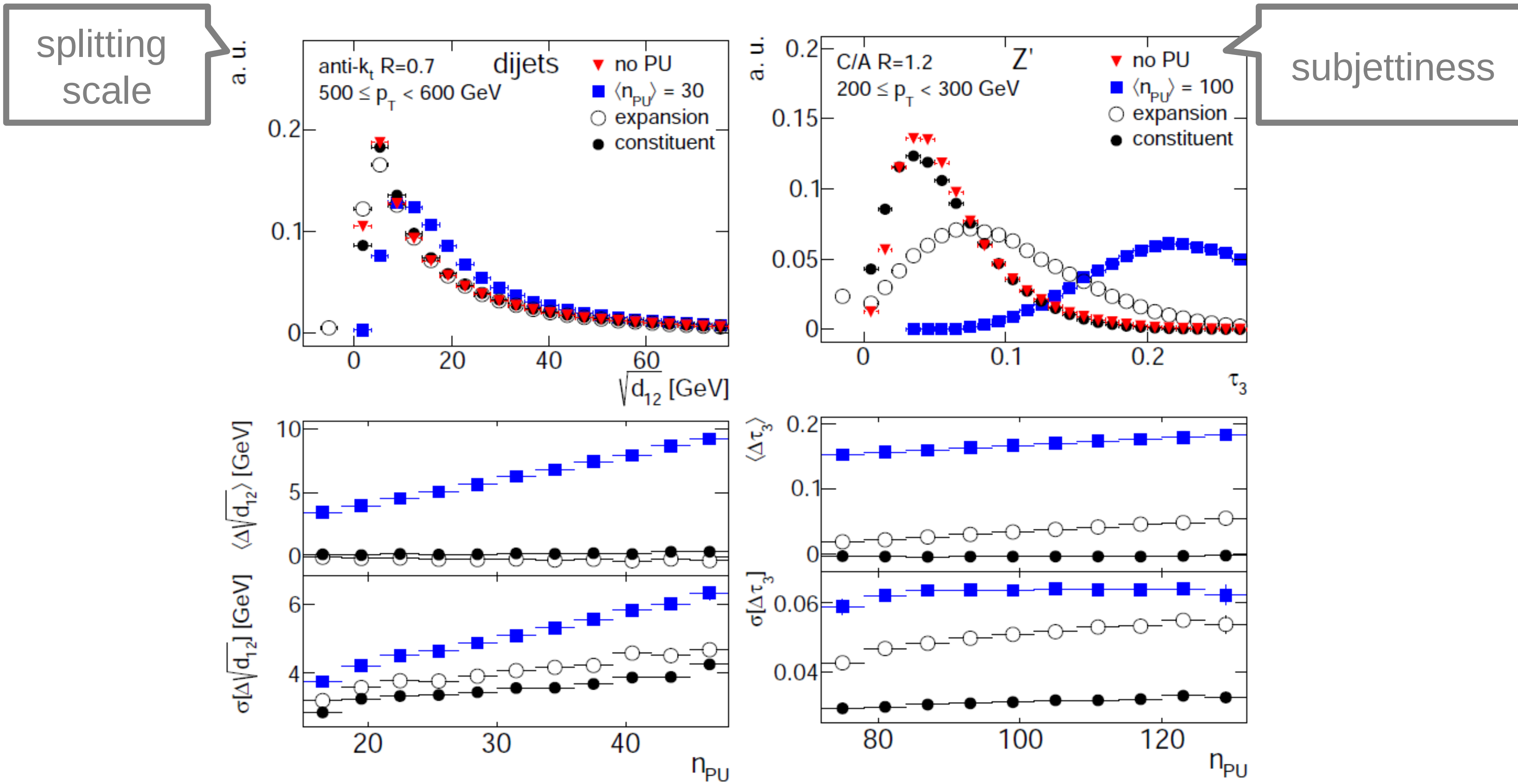
mean jet mass = corrected – truth

mass resolution



Significant improvement over area based subtraction

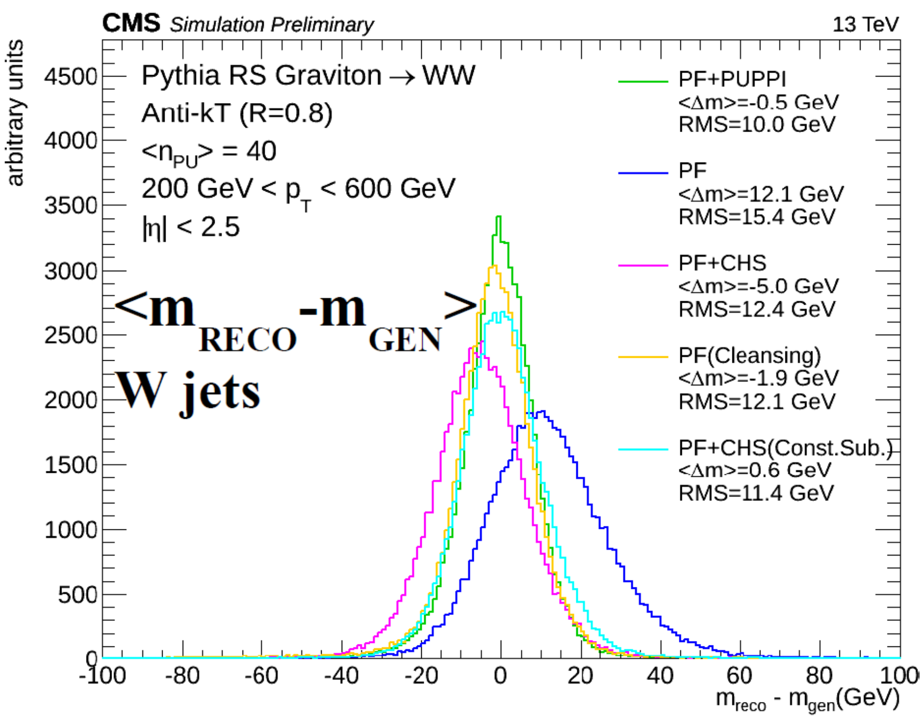
Performance – jet shape variables

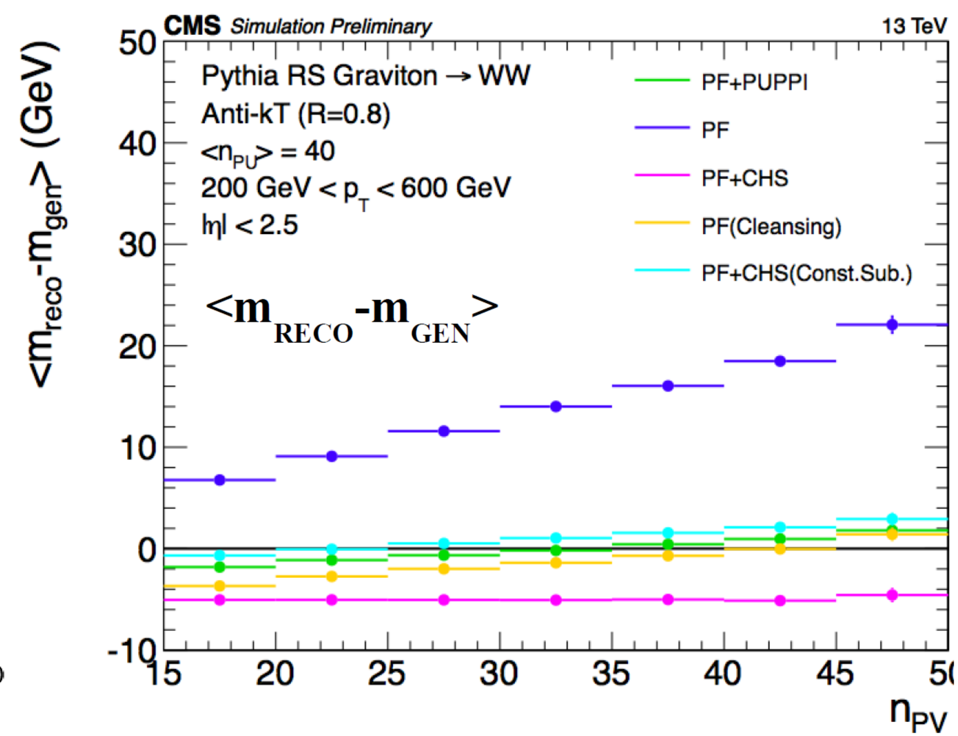


- Examples of variables often used to discriminate highly boosted objects.
- Significant improvement comparing to the shape expansion method.

Performance – full detector simulation

Viola Sordini, BOOST 2014





- Tests using the full detector simulation by CMS.
- Very good performance of both PUPPI and Constituent subtraction.
- PUPPI gives slightly better resolutions than Constituent subtraction.

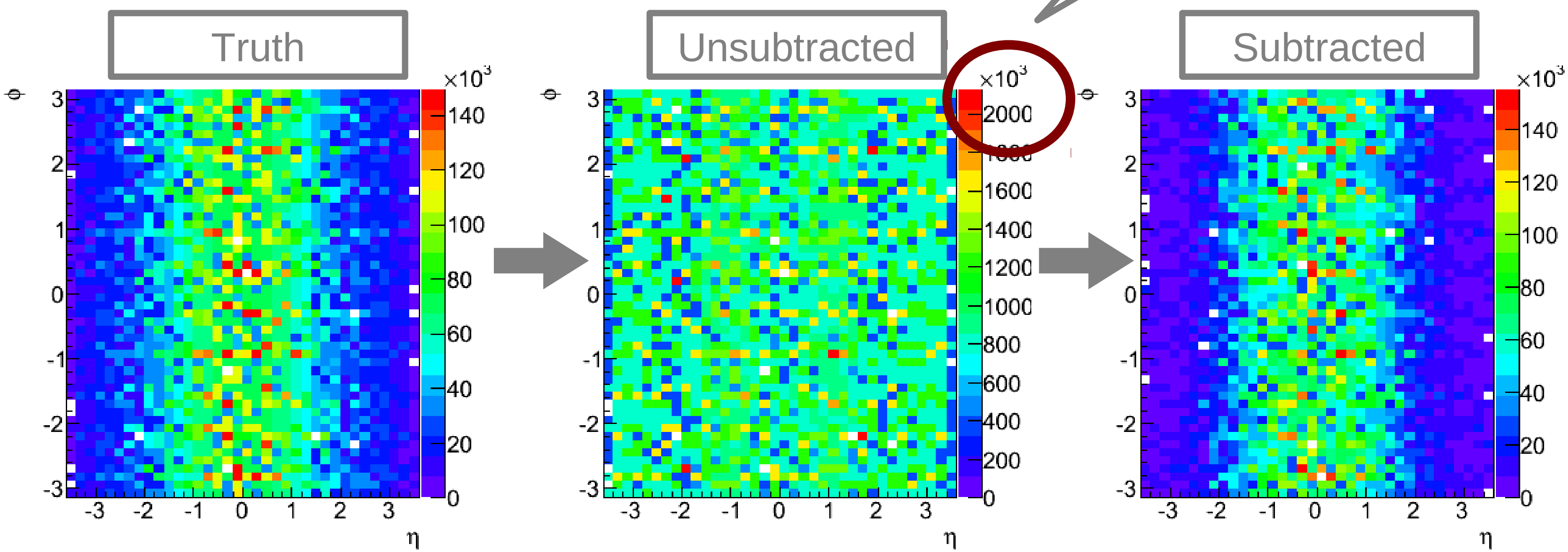
Full event subtraction

Algorithm allows for full event subtraction

–> may allow event shape studies

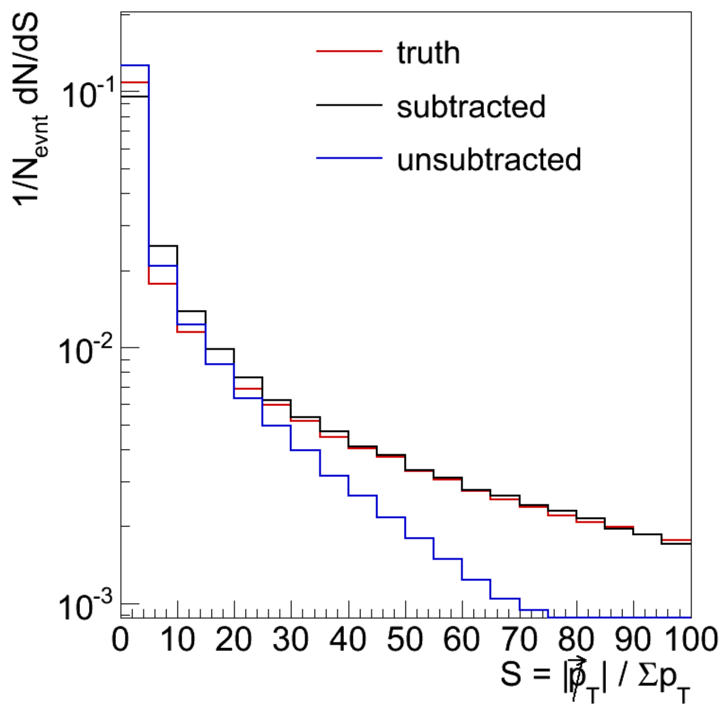
–> may improve missing-pt reconstruction

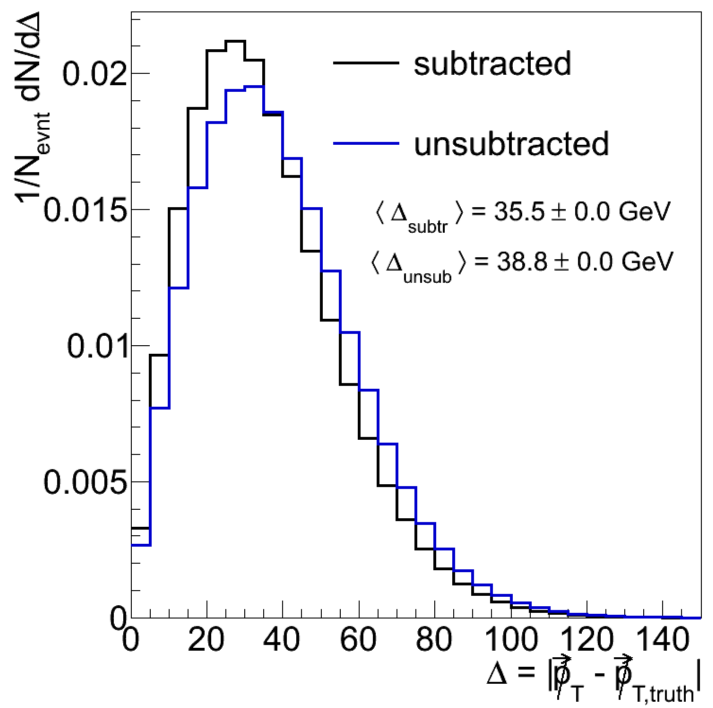
Number of pseudo-towers with pt>0



Missing-pt significance

- Trivial missing-pt reconstruction on subtracted events tested in Z' events with mu=100, $\vec{\not p}_T = (\Sigma\, p_T \cos\phi, \Sigma\, p_T \sin\phi)$
- Two kinds of missing-pt significance variables estimated. Subtracted events significantly better agreement with truth than unsubtraced.
- Further improvements possible if subtracted objects enter estimates.





Conclusions

- Pileup/UE subtraction is a very important problem in pp/HI collisions!
- Several new methods for the pileup subtraction on the market. Methods should be tested side by side. Ideas should be shared.
- Constituent subtraction has a very good ability to remove the pileup (or HI underlying event).
- Constituent subtraction is probably the most generic method.
- It is implemented in FastJet Contrib and easy to plug in to the reconstruction chain.

Backup

Two basic methods of grooming

- Trimming
  - reconstruct jets with large radii (e.g. R=1.0) and recluster to jets with smaller radii (e.g. R_{trim}=0.2)
  - accept only if their f = $p_T/p_{T,orig}$ grater than e.g. 0.03
  - recombine surviving sub-jets into a groomed jet
- Pruning
  - run Cambridge-Aachen algorithm on jet constituents
  - during clustering of old-constituents *i* and *j* the softer constituent is removed if

$$z_{ij} = \frac{\min(p_{T,i}, p_{T,j})}{p_{T,i} + p_{T,j}} < z_{\mathrm{cut}}$$

$$\Delta R_{i,j} > D_{\mathrm{cut}} = \frac{2\ r_{\mathrm{cut}}\ m_{\mathrm{orig}}}{p_{T,\mathrm{orig}}}$$

  - typical size of parameters: z_{cut} = 0.1, r_{cut} = 0.5
- Other methods: filtering, mass drop.

Subtracting mass

$$\rho = \mathrm{median}_{\mathrm{patches}}\left\{\frac{p_{\mathrm{T}\mathrm{patch}}}{A_{\mathrm{patch}}}\right\}$$

$$\rho_m = \mathrm{median}_{\mathrm{patches}}\left\{\frac{m_{\delta\mathrm{patch}}}{A_{\mathrm{patch}}}\right\} \qquad m_\delta = \sqrt{m^2 + p_{\mathrm{T}}^2} - p_{\mathrm{T}}$$

$$p_{\mathrm{T}}^g = A_g \cdot \rho,$$
$$m_\delta^g = A_g \cdot \rho_m$$

$$\begin{aligned} \text{If } p_{\mathrm{T}i} \geq p_{\mathrm{T}k}^g : \quad & p_{\mathrm{T}i} \longrightarrow p_{\mathrm{T}i} - p_{\mathrm{T}k}^g, \\ & p_{\mathrm{T}k}^g \longrightarrow 0; \\ \text{otherwise:} \quad & p_{\mathrm{T}i} \longrightarrow 0, \\ & p_{\mathrm{T}k}^g \longrightarrow p_{\mathrm{T}k}^g - p_{\mathrm{T}i}. \end{aligned}$$

$$\begin{aligned} \text{If } m_{\delta i} \geq m_{\delta k}^g : \quad & m_{\delta i} \longrightarrow m_{\delta i} - m_{\delta k}^g, \\ & m_{\delta k}^g \longrightarrow 0; \\ \text{otherwise:} \quad & m_{\delta i} \longrightarrow 0, \\ & m_{\delta k}^g \longrightarrow m_{\delta k}^g - m_{\delta i}. \end{aligned}$$

Constituent subtraction

$$p_{\mathrm{corr}}^\mu = [p^x - \rho A^x, p^y - \rho A^y, p^z - (\rho + \rho_m) A^y, E - (\rho + \rho_m) A^E]$$

$$A_\mu = \frac{A_g}{p_{T,g}} \sum_{\substack{k \in \mathrm{ghosts} \\ \mathrm{in\ jet}}} p_{g,k}^\mu$$

4-vector area-median subtraction

Jet shapes

- N-subjettiness:

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T}k} \cdot \min(\Delta R_{1k}, \Delta R_{2k}, ..., \Delta R_{Nk})$$

$$d_0 \equiv \sum_k p_{\mathrm{T}k} \cdot R$$

  - reclustering the jet with kt algorithm, requiring to find N subjets
  - distances between subjet axes and constituents
  - also ratios of t_N can be used
- kt splitting scale

$$\sqrt{d_{12}} = \min(p_{\mathrm{T}}^1, p_{\mathrm{T}}^2) \cdot \Delta R_{12}$$

  - transverse momenta of subjets and the distance between them