

NpC and SoftKiller pileup removal

Matteo Cacciari (LPTHE Paris), Gavin Salam (CERN), Gr  gory Soyez (Saclay)

arXiv:1404.7353 and arXiv:1407.0408

Recent progress on pileup removal

A lot of recent activity:

- ▶ CMS Voronoi method ([Lai, unpubl.](#))
- ▶ Cleansing ([Krohn, Schwartz, Low, Wang, 1309.4777](#))
- ▶ corrJVF ([ATLAS-PHYS-PUB-2014-001](#))
- ▶ Constituent Subtraction ([Berta, Spousta, Miller, Leitner, 1403.3108](#))
- ▶ NpC ([MC, Salam, Soyez, 1404.7353](#))
- ▶ PUPPI ([Bertolini, Harris, Low, Tran, 1407.6013](#))
- ▶ SoftKiller ([MC, Salam, Soyez, 1407.0408](#))
- ▶ ...

Various methods reviewed and compared at the CERN pileup workshop in May
<https://indico.cern.ch/event/306155/>

Pileup subtraction methods

Full jet/**Observable** level

- ▶ Determination of *susceptibility to contamination* of each specific observable needed
- ▶ Basic example: transverse momentum
 $\mathbf{p}_t^{\text{sub}} = \mathbf{p}_t^{\text{raw}} - \rho \mathbf{A}$ ([MC, Salam 0707.1378](#))
- ▶ Other examples:
 - ▶ Analytical calculations of susceptibility for selected jet shapes ([Sapeta et al. 1009.11143](#), [Alon et al. 1101.3002](#))
 - ▶ Moments of jet fragmentation functions ([MC, Quiroga, Salam, Soyez, 1209.6086](#))
 - ▶ Generic (numerical) approach to susceptibility determination for any shape ([Soyez et al, 1211.2811](#))

Subjet/particle level

Pileup subtraction methods

Full jet/**Observable** level

- ▶ Determination of *susceptibility to contamination* of each specific observable needed
- ▶ Basic example: transverse momentum
 $\mathbf{p}_t^{\text{sub}} = \mathbf{p}_t^{\text{raw}} - \rho \mathbf{A}$ ([MC, Salam 0707.1378](#))
- ▶ Other examples:
 - ▶ Analytical calculations of susceptibility for selected jet shapes ([Sapeta et al. 1009.11143](#), [Alon et al. 1101.3002](#))
 - ▶ Moments of jet fragmentation functions ([MC, Quiroga, Salam, Soyez, 1209.6086](#))
 - ▶ Generic (numerical) approach to susceptibility determination for any shape ([Soyez et al, 1211.2811](#))

Subjet/**particle** level

- ▶ The event is modified before calculating observables (jets, shapes, etc). Corrections applied to subjets or even to particles
- ▶ Examples:
 - ▶ Cleansing ([Krohn, Schwartz, Low, Wang, 1309.4777](#))
 - ▶ corrJVF ([ATLAS-PHYS-PUB-2014-001](#))
 - ▶ NpC ([MC, Salam, Soyez, 1404.7353](#))
 - ▶ CMS Voronoi method ([Lai, unpubl.](#))
 - ▶ Constituent Subtraction ([Berta, Spousta, Miller, Leitner, 1403.3108](#))
 - ▶ PUPPI ([Bertolini, Harris, Low, Tran, 1407.6013](#))
 - ▶ SoftKiller ([MC, Salam, Soyez, 1407.0408](#))
 - ▶ ...

Pileup subtraction methods

Full jet/**Observable** level

Subjet/particle level

► Pros:

- ▶ Subtraction is unbiased by construction
- ▶ Not too sensitive to detector effects (works at the jet/subjet level)

► Cons:

- ▶ Need to cluster (e.g. to calculate areas), hence time-consuming

Pileup subtraction methods

Full jet/**Observable** level

▶ Pros:

- ▶ Subtraction is unbiased by construction
- ▶ Not too sensitive to detector effects (works at the jet/subjet level)

▶ Cons:

- ▶ Need to cluster (e.g. to calculate areas), hence time-consuming

Subjet/particle level

▶ Pros:

- ▶ Often no need to cluster
- ▶ Dispersion usually reduced
- ▶ If left with a ‘subtracted’ event, one can then calculate any observable

▶ Cons:

- ▶ Potentially sensitive to detector effects
- ▶ Potentially biased (but can usually be tuned)

- ▶ **Working hypothesis:** pileup energy flow is distributed sufficiently uniformly over the event
- ▶ Estimate **pileup transverse momentum density ρ** , using measurements of energy flow in patches of given size
 - ▶ Possibly rescale ρ as a function of rapidity and azimuth
- ▶ Calculate **area A_μ** of each jet
- ▶ Subtract pileup contamination using

$$p_\mu^{\text{jet,sub}} = p_\mu^{\text{jet,full}} - \rho A_\mu$$

(This method can be adapted to jet shapes)

Neutral proportional to Charged

► Working hypotheses:

arXiv:1404.7353

1. one can **detect all charged** particles from pileup,
and therefore **measure** $p_\mu^{\text{jet,chg-PU}}$
2. the momentum from the unseen **neutral** component of
pileup is **proportional** to the measured **charged** one, i.e.
there exists a **fixed charged fraction** $\gamma_0 = p^{\text{chg-PU}} / p^{\text{PU}}$

► Then, two options:

- Use **full** event, and subtract pileup as

$$p_\mu^{\text{jet,sub}} = p_\mu^{\text{jet,full}} - \frac{1}{\gamma_0} p_\mu^{\text{jet,chg-PU}}$$

- Use **CHS** (= charged hadron subtracted) event, i.e. without charged particles from pileup (technically, scaled by $\epsilon \ll 1$) and subtract as

$$p_\mu^{\text{jet,sub}} = p_\mu^{\text{jet,CHS}} - (1 - \gamma_0) \frac{1}{\gamma_0} \frac{p_\mu^{\text{jet,rescaled-chg-PU}}}{\epsilon}$$

When can one expect things to work well?

- ▶ For **area-median**, if point-to-point pileup energy-flow fluctuations are moderate (since ρ is estimated globally)
- ▶ For **NpC**, if energy flows from neutral and charged pileup particles are really spatially well correlated

Which one wins?

area-median v. NpC

Check how well each method estimates the neutral component of pileup transverse momentum

I. **Estimate p_t^{ntr}** in a patch of radius R using

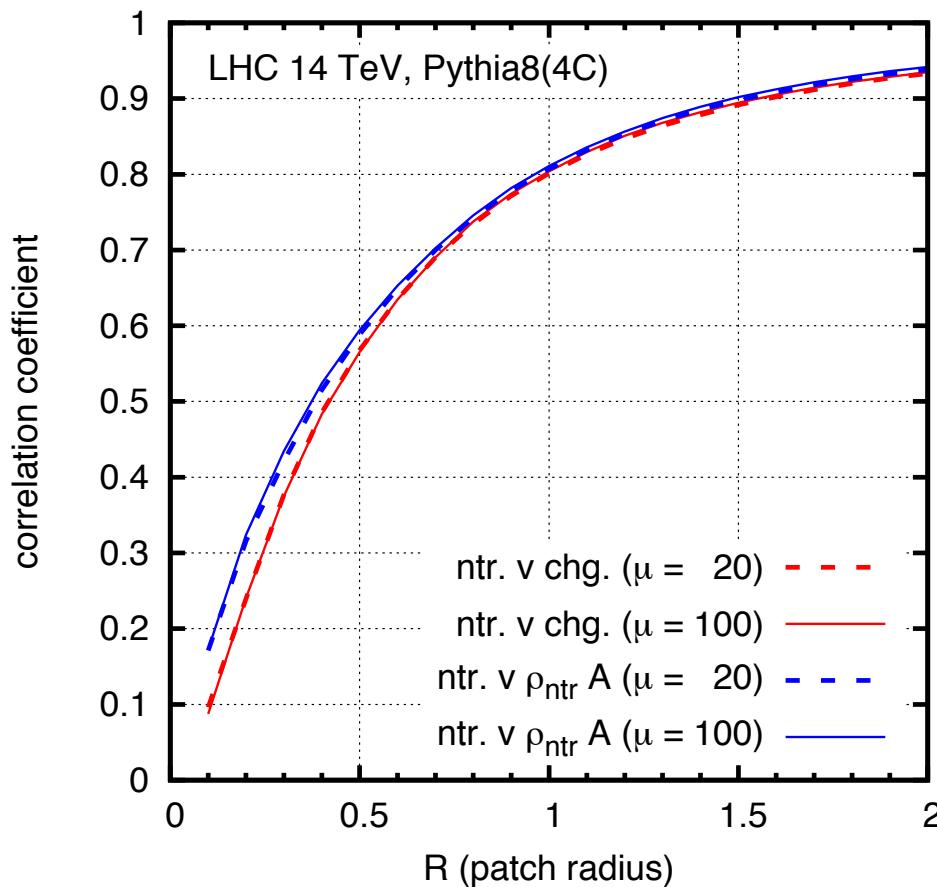
- ▶ either $\rho^{ntr} A$
- ▶ or $p_t^{chg} (1 - \gamma_0) / \gamma_0$

2. **Determine quality of estimation** by looking

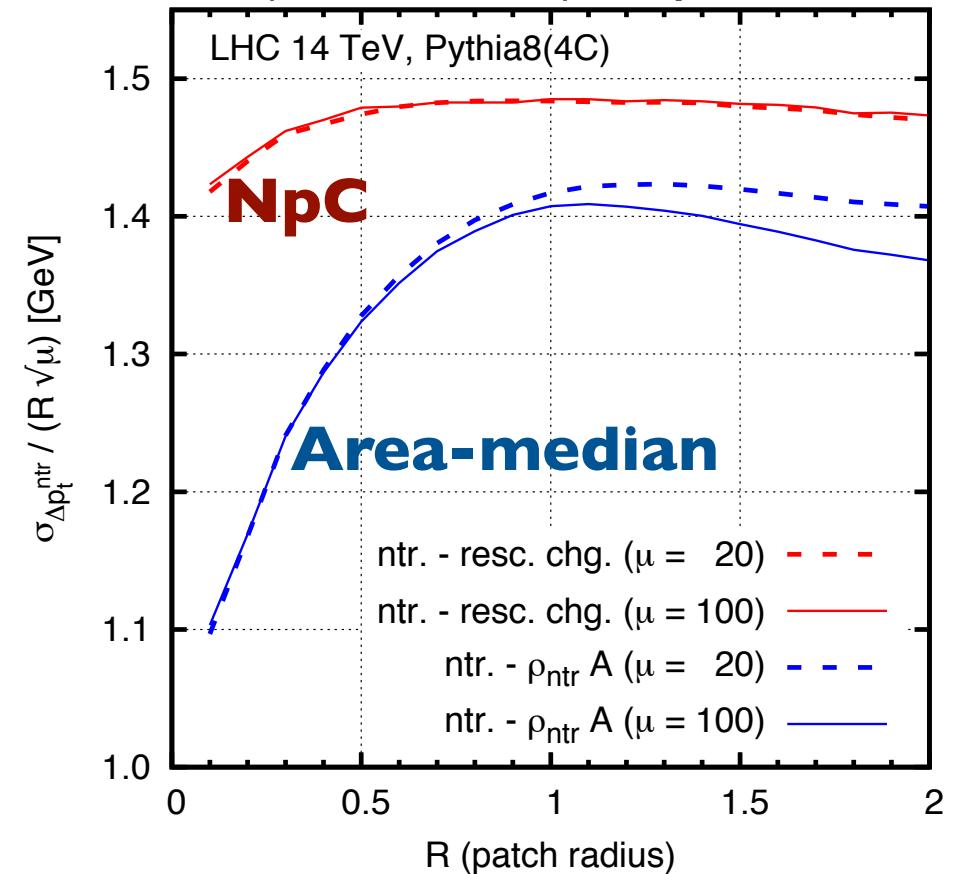
- ▶ at the correlation coefficient between p_t^{ntr} and $\rho^{ntr} A$ or p_t^{chg}
- ▶ at the dispersion of the “misestimations”: $p_t^{ntr} - \rho^{ntr} A$ or $p_t^{ntr} - p_t^{chg} (1 - \gamma_0) / \gamma_0$

area-median v. NpC

Correlation coefficient



(normalized) dispersion

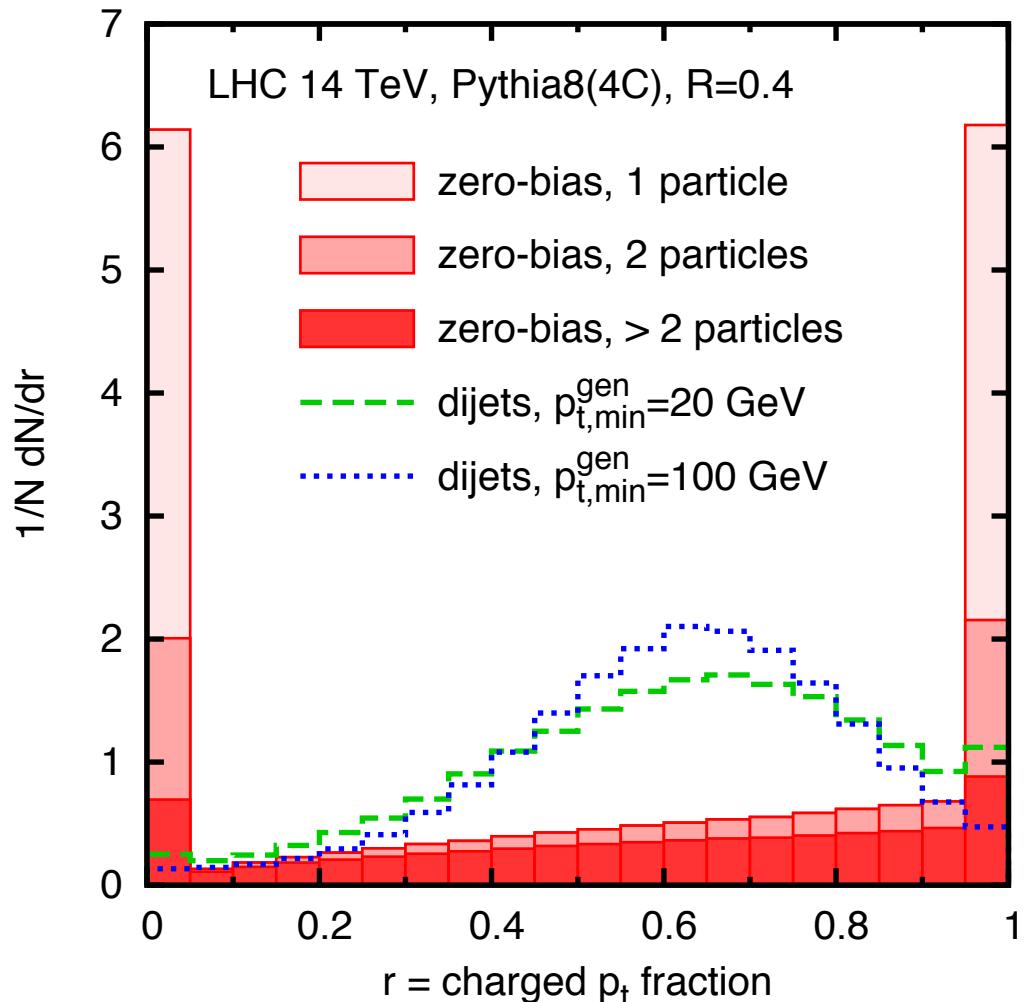


NpC is no better than area-median at estimating the neutral component of pileup p_t^{ntr} in a patch

In fact, area-median is slightly better at all values of R , and especially at small R (< 0.5)

Why is NpC no better?

Short answer: because local correlation between neutral and charged is not that great for pileup



Charged pt fraction in

- ▶ a patch of pileup of radius 0.4
- ▶ anti- kt jet with $R=0.4$

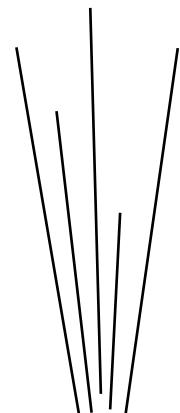
Most of the time, the pileup patch contains only a few charged or a few neutral particles from a given PU vertex, and the charged fraction is **not** peaked around a single value as in the case of the hard jets

Why is NpC no better?

Marginally longer answer: because decays of pileup particles tend to take place at large angle

hard jet

$$p_t \gg \Lambda_{\text{QCD}}$$



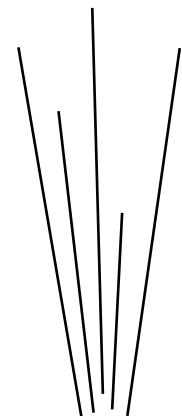
hadronisation

Charged and neutral energy tend to go in the same direction

Why is NpC no better?

Marginally longer answer: because decays of pileup particles tend to take place at large angle

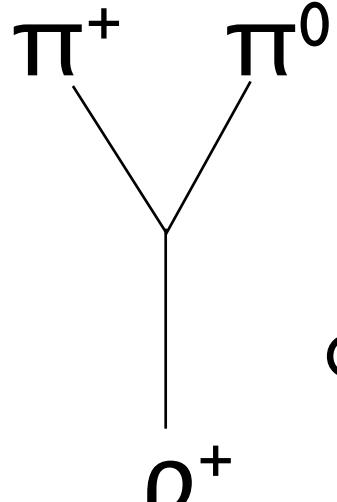
hard jet
 $p_t \gg \Lambda_{\text{QCD}}$



hadronisation

Charged and neutral energy tend to go in the same direction

pileup
 $p_t \simeq 0.5 - 2 \text{ GeV}$



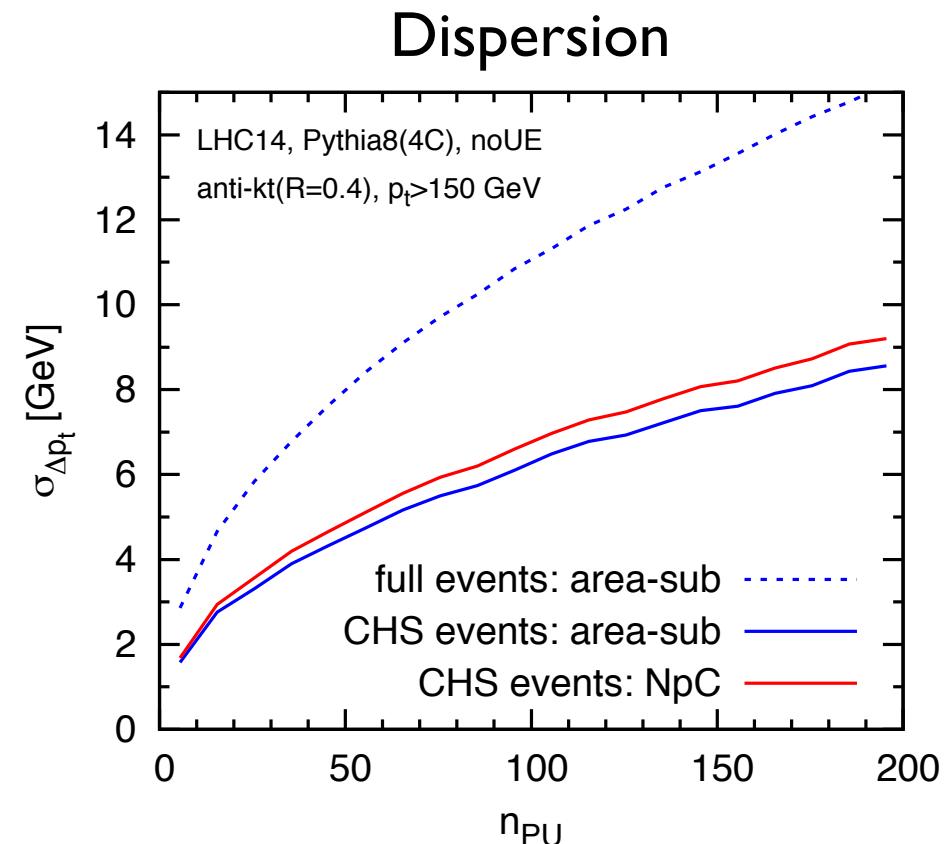
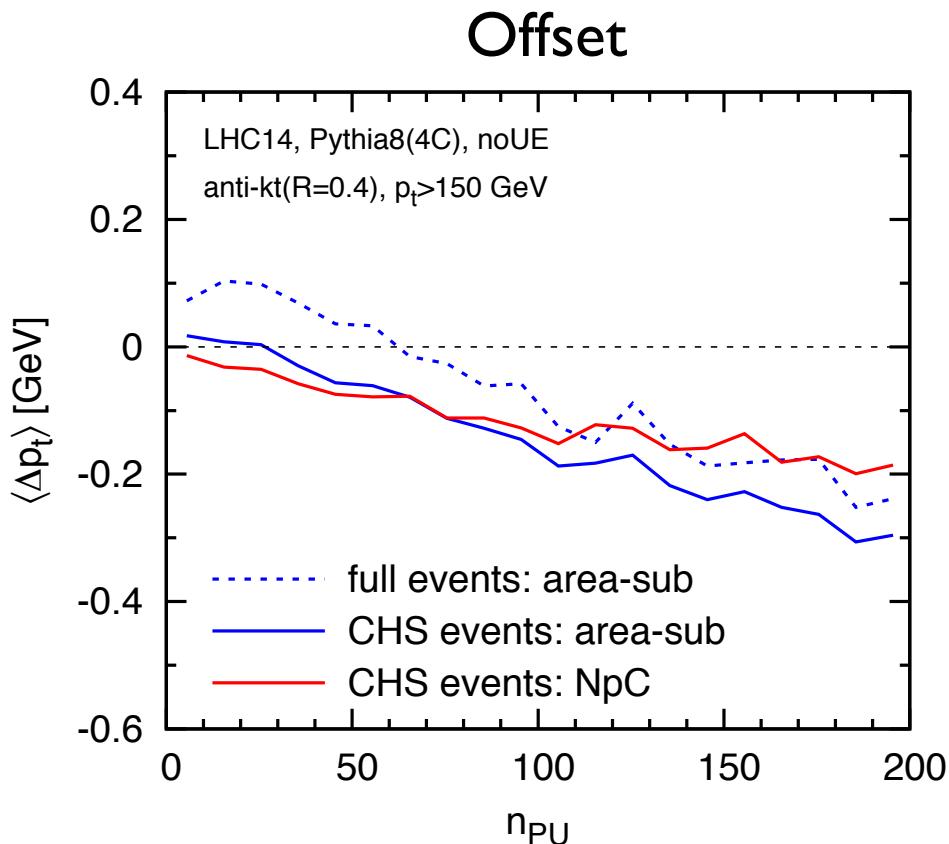
one step of
hadronisation

Opening angle $\sim 2m_\rho/p_{t,\rho} \sim 1$

Emissions at large angles break
local correlation

Performance of NpC

Pileup subtraction in dijet production Observable: jet p_t

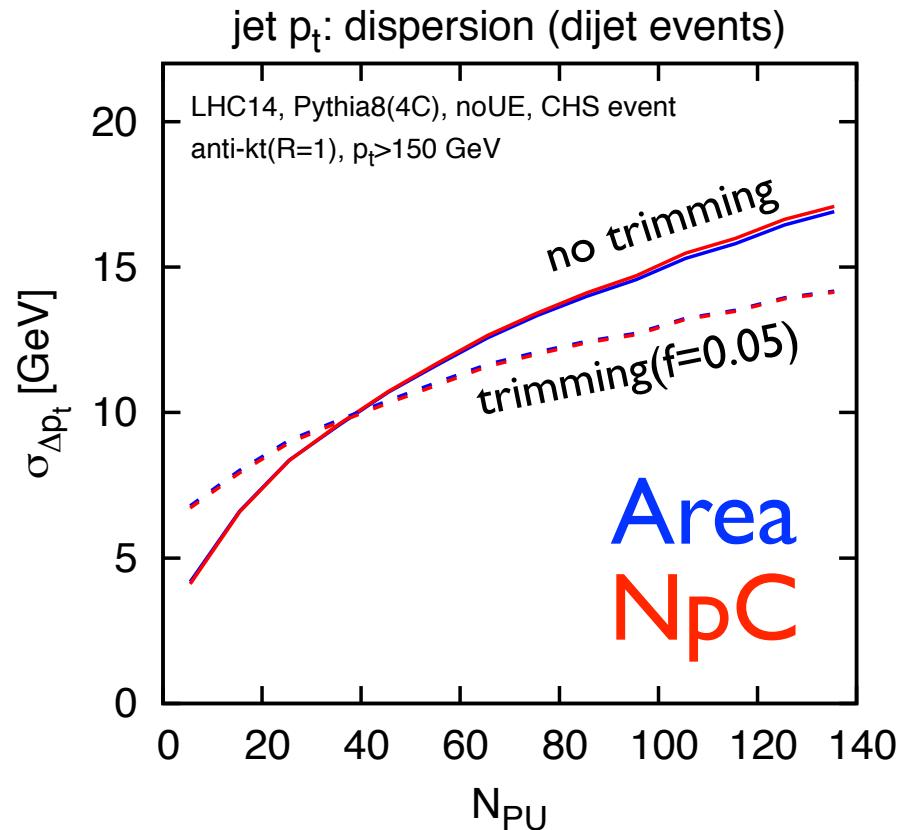
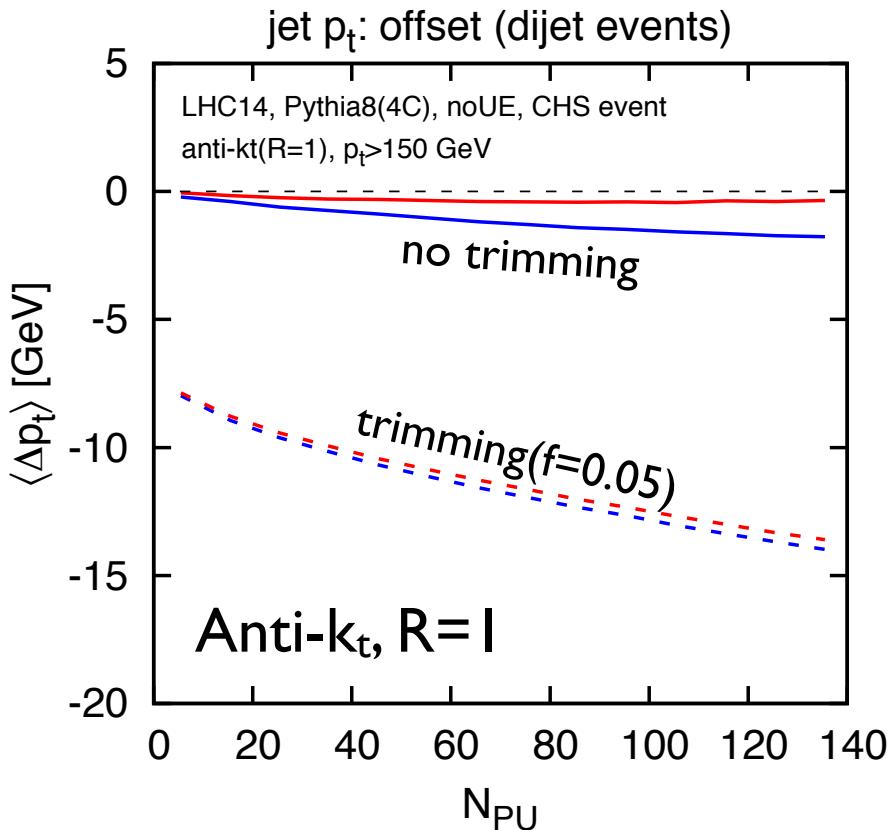


Offset limited to
a few hundred MeV

As expected,
 $\sigma^{CHS,NpC} \sim \sigma^{CHS,area} < \sigma^{full,area}$

Addition of grooming

Study effect of adding trimming ($f=0.05$)



As expected, trimming reduces the dispersion (only some subjets are retained) at the price of increasing the offset (it cuts into hard perturbative radiation)

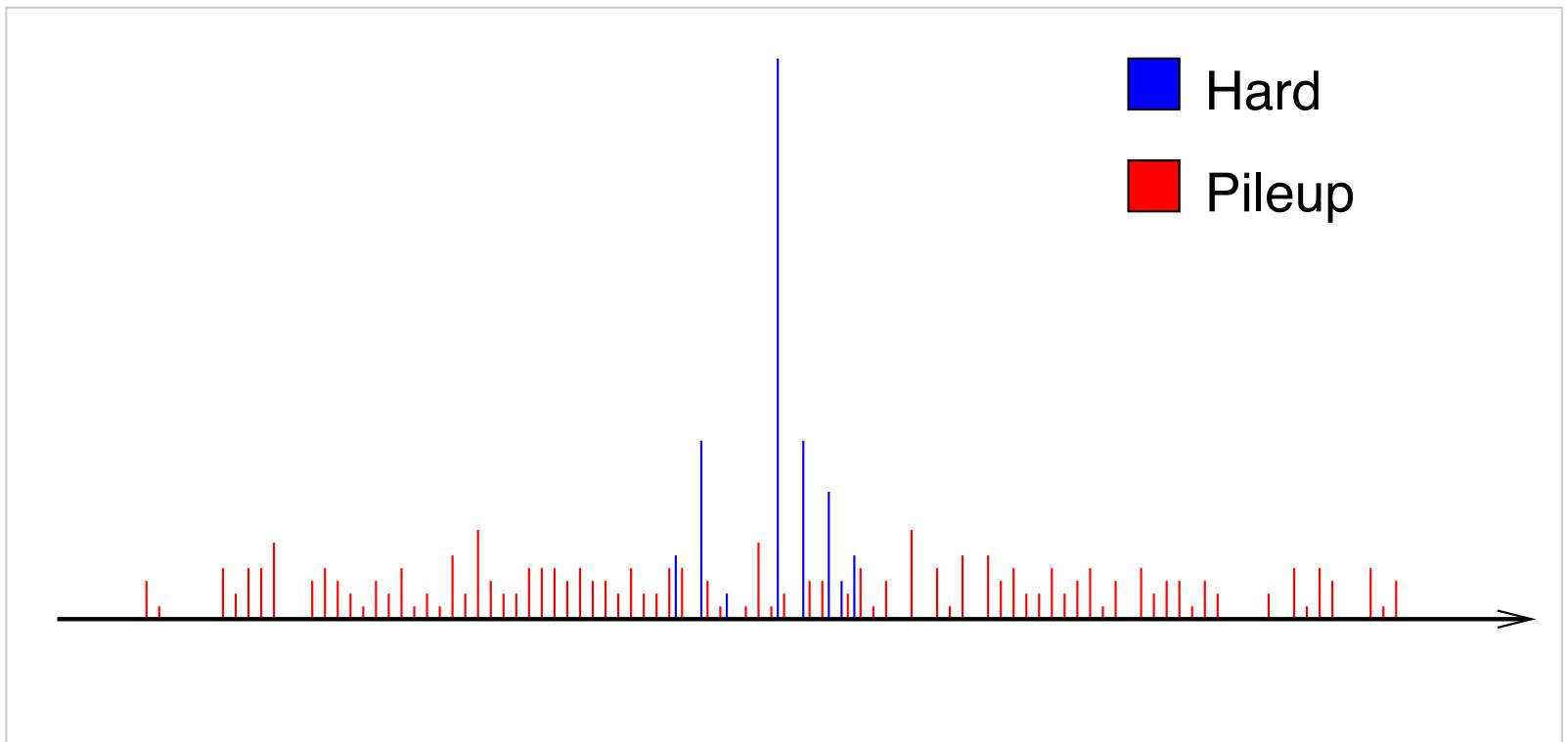
Take-home message

► **Area-median** and **NpC** perform similarly

- Perhaps contrary to intuition, NpC does not work better, because large angle emissions of low-pt pileup particle tend to destroy local correlation between neutral and charged particles
- The dispersion of both methods scales with $\sqrt{N_{PU}}$

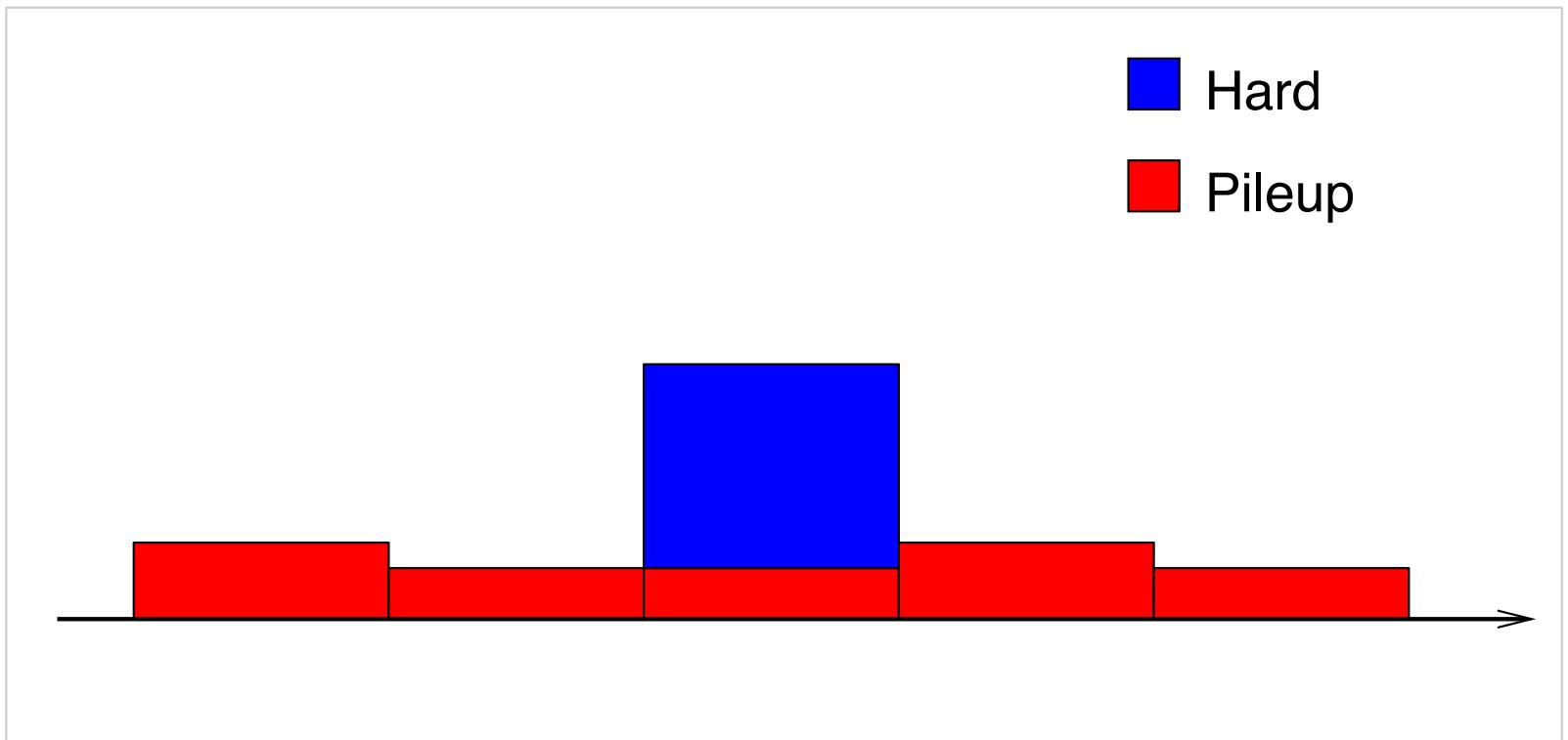
- ▶ A very simple, and fast, event-level pileup removal method
- ▶ Essentially a p_T cut with a dynamically-determined threshold

An event: particle level



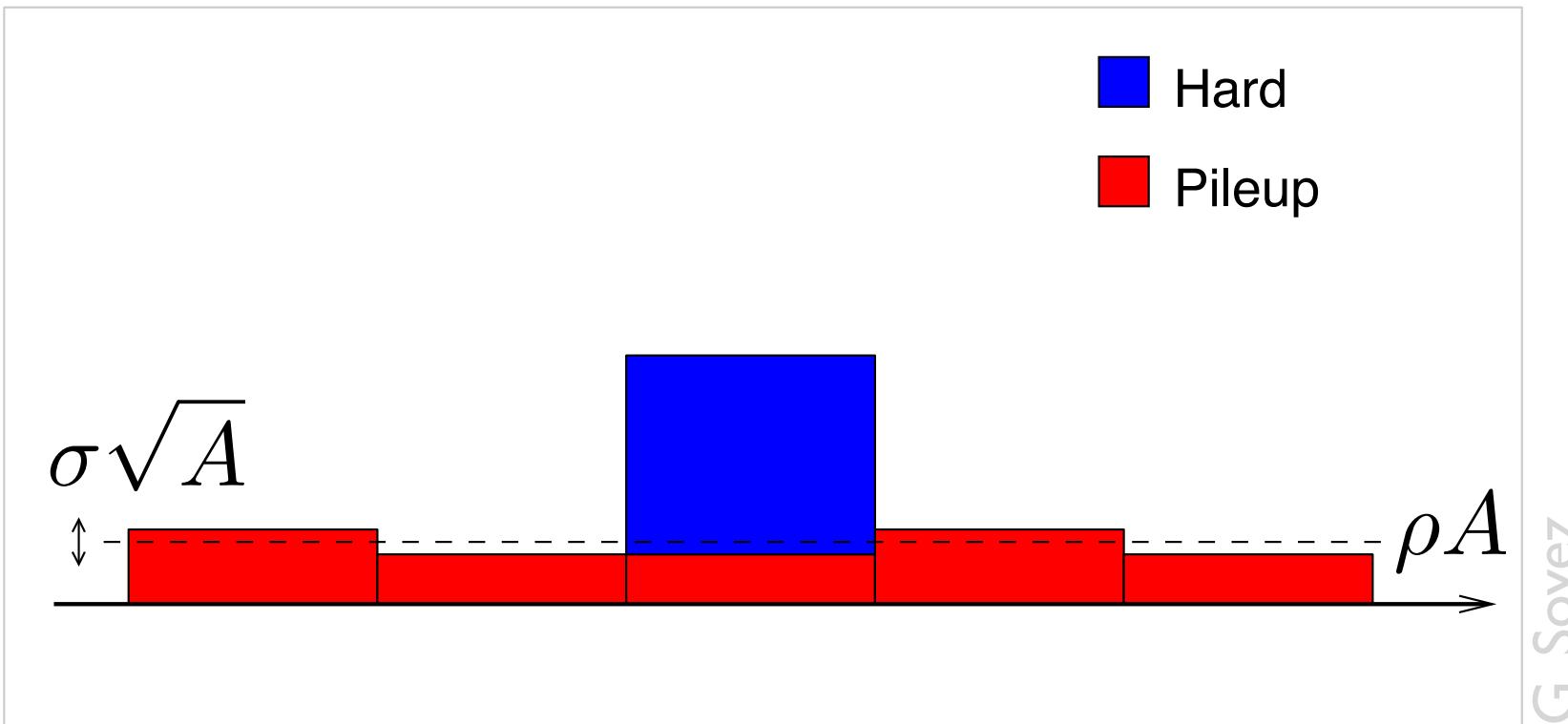
G. Soyez

An event: jet level



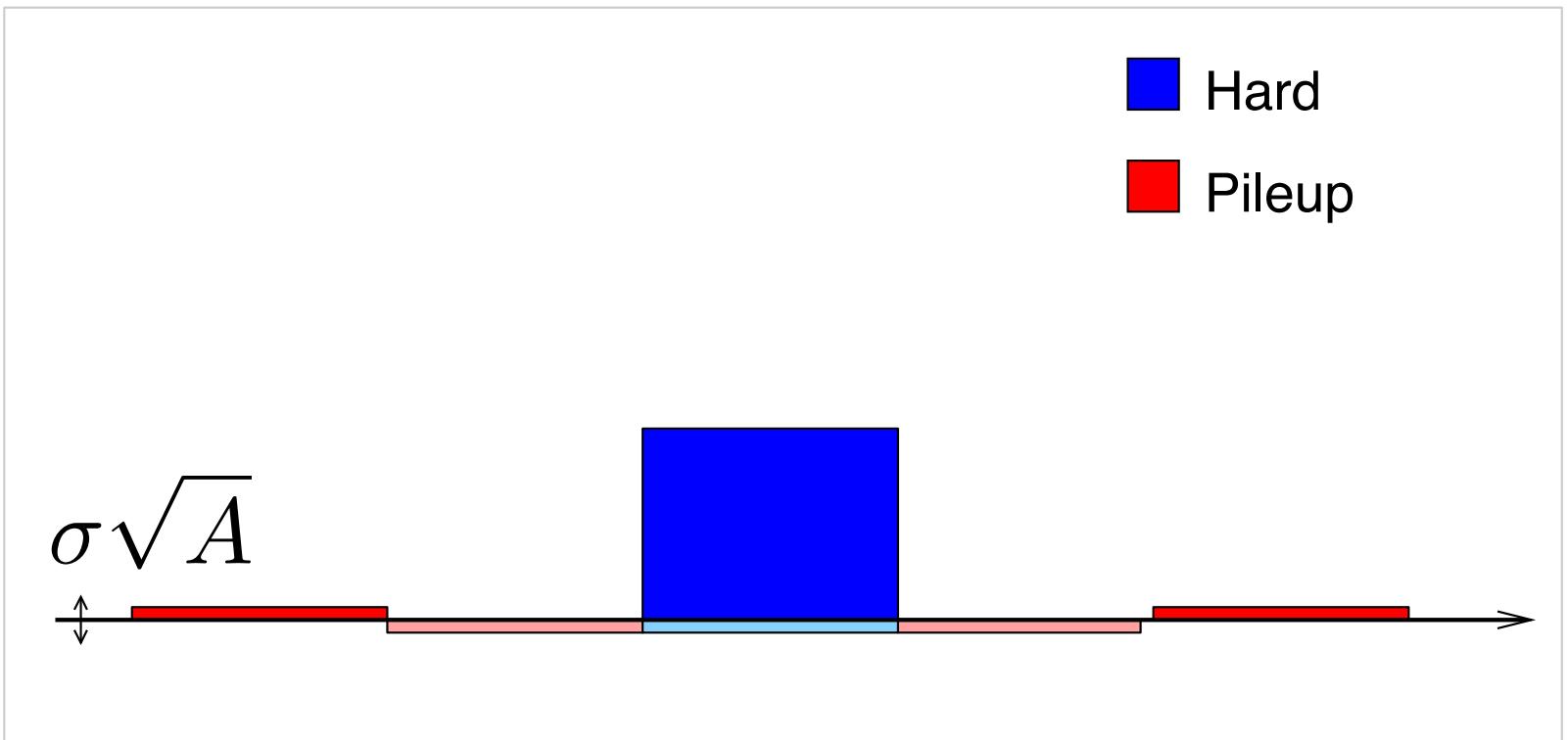
G. Soyez

Background determination

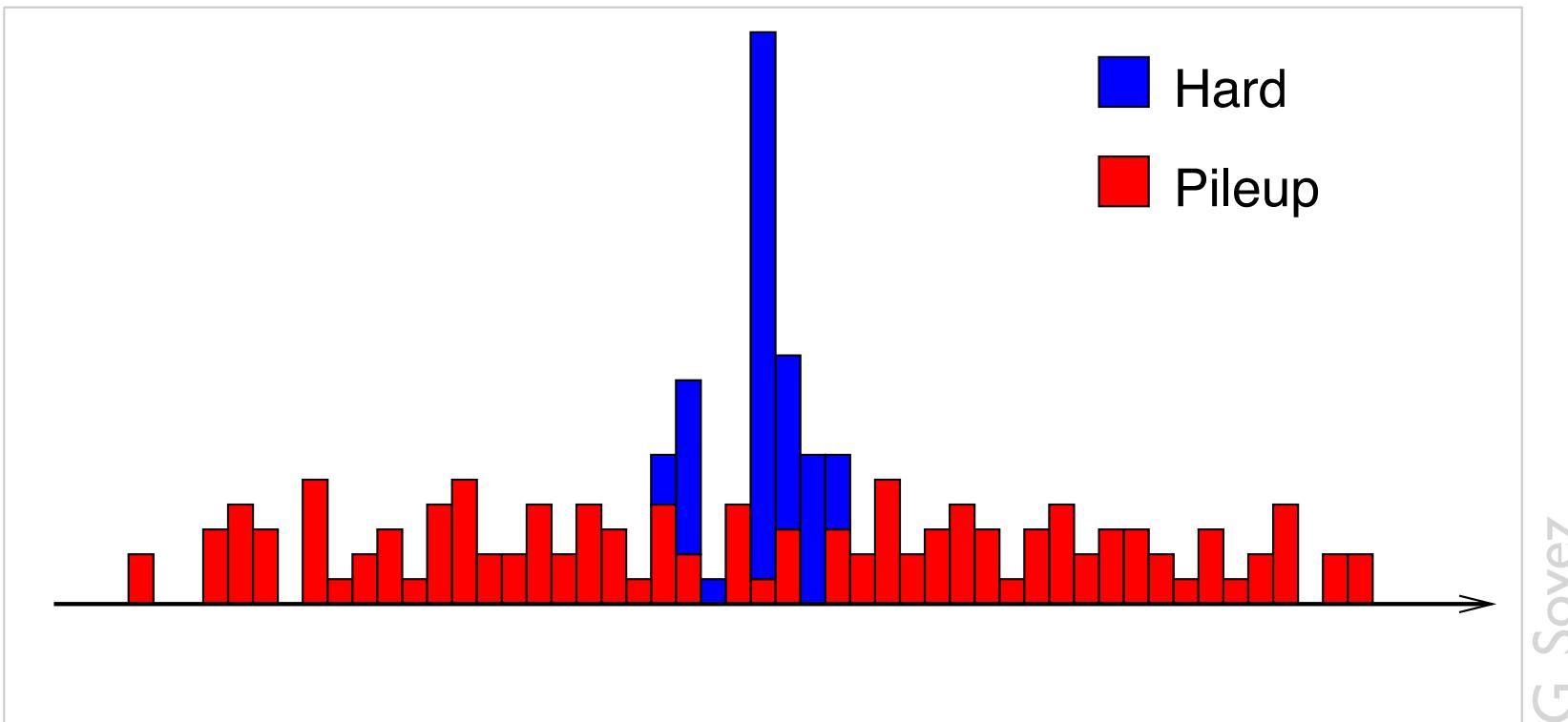


G. Soyez

Background subtraction

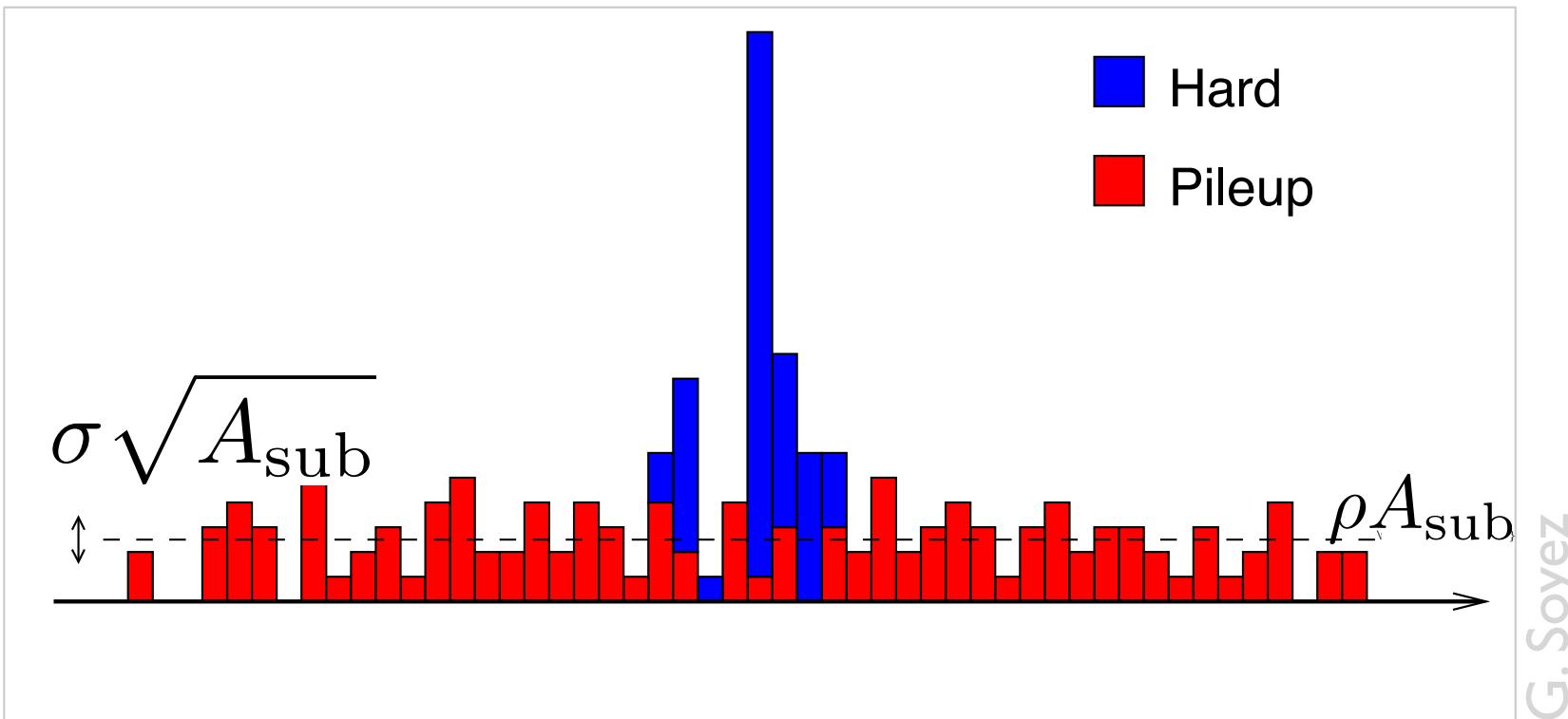


Use subjects instead



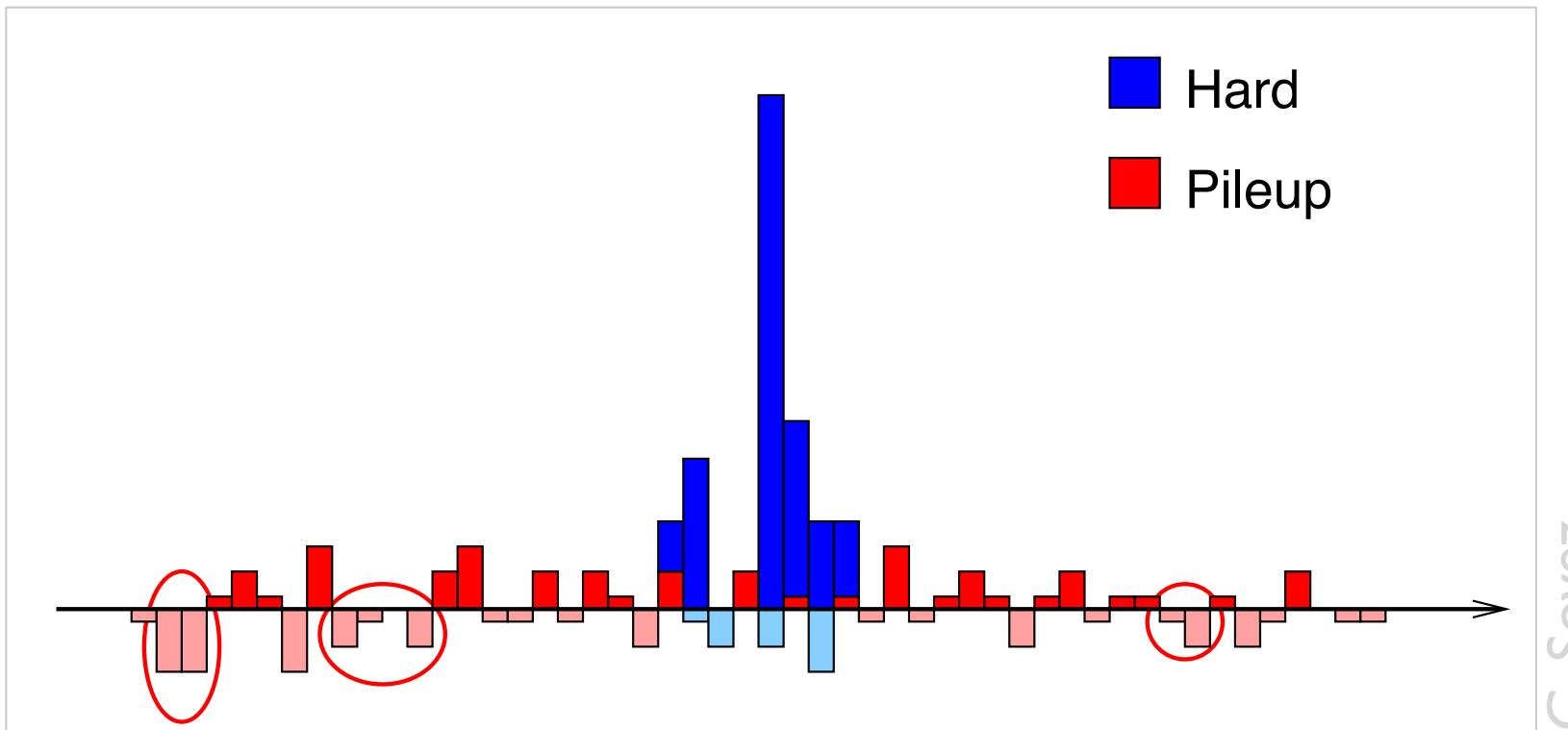
G. Soyez

Use subjects instead



G. Soyez

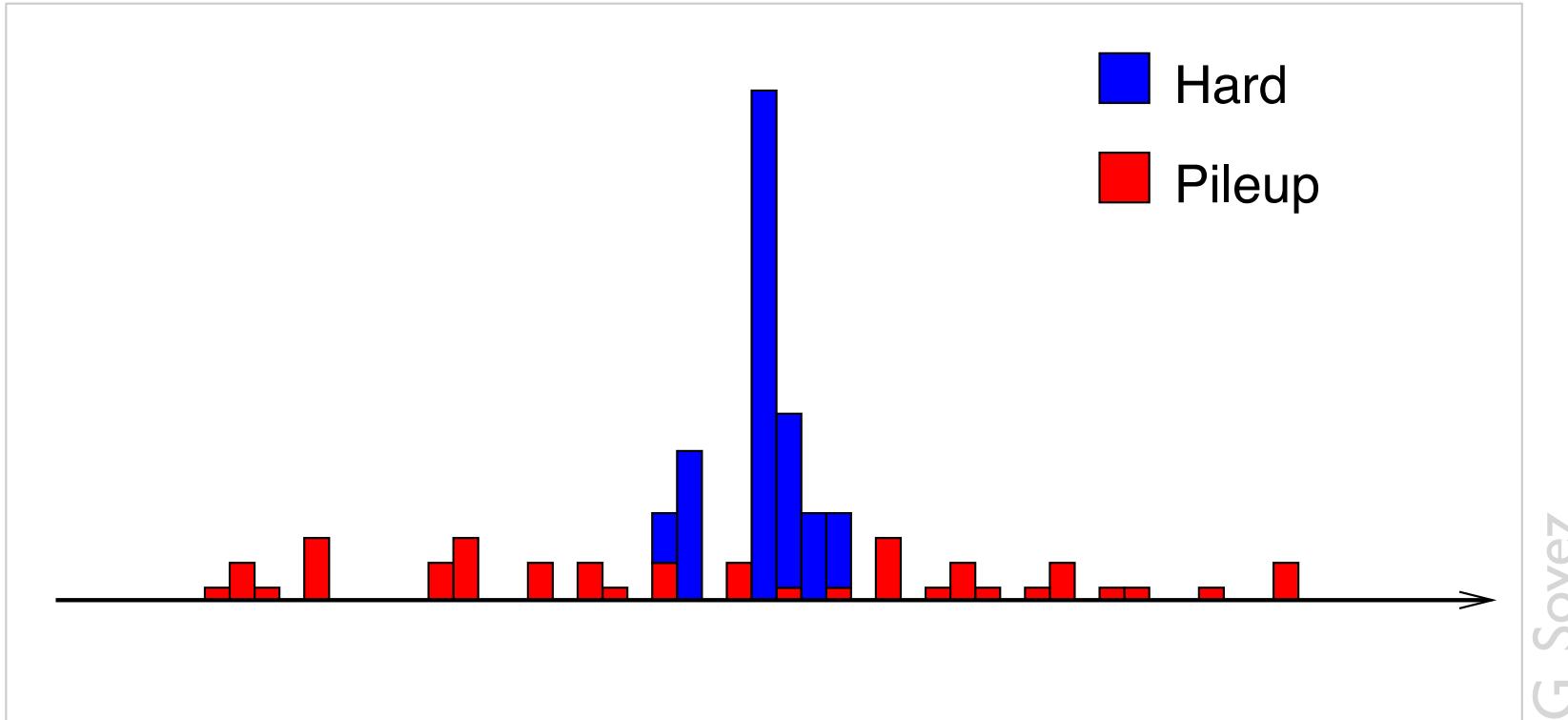
Use subjets instead



G. Soyez

Same average (~ 0) and dispersion on the scale of the full jet,
but some subjets are negative

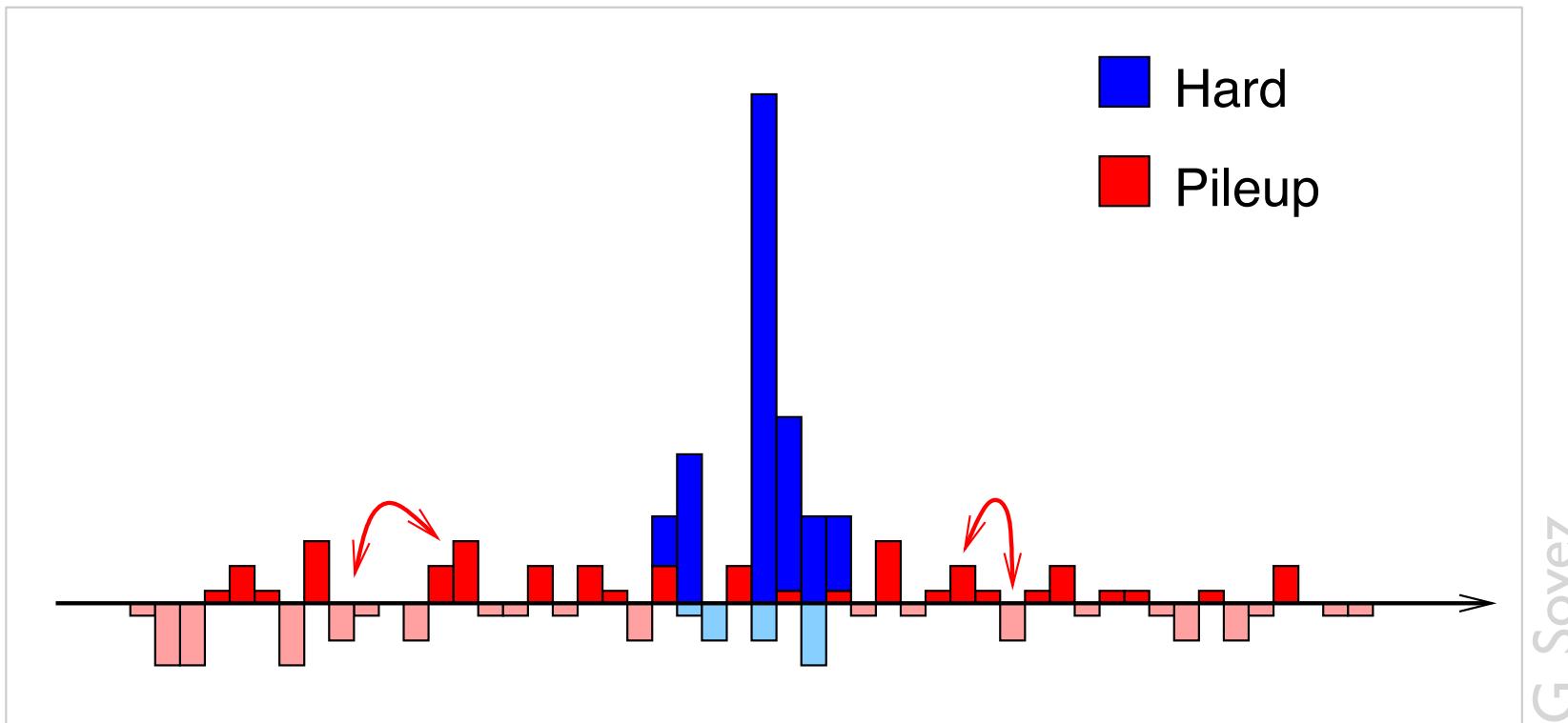
Use subjects instead



G. Soyez

Most naive noise reduction approach: cut away the negative part.
Biased.

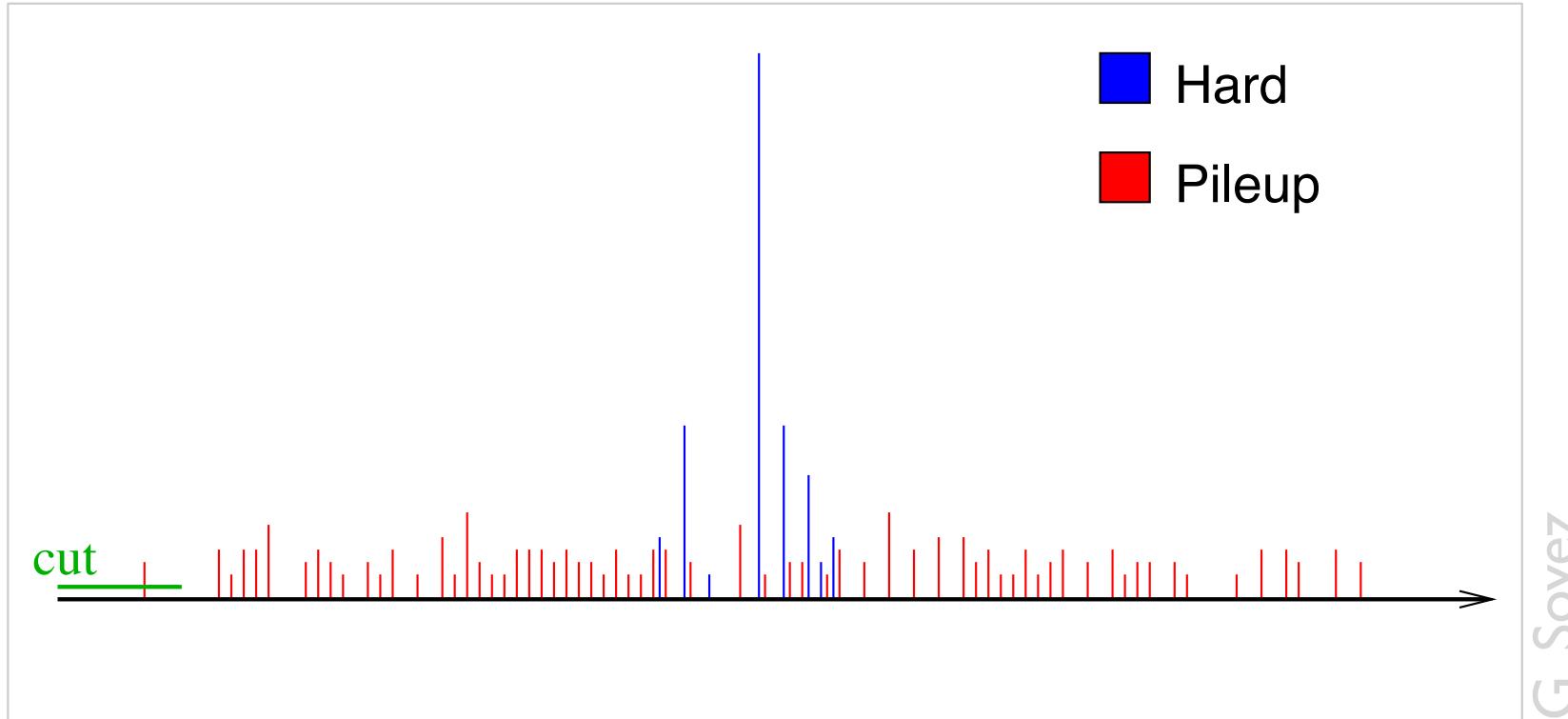
Use subjects instead



G. Soyez

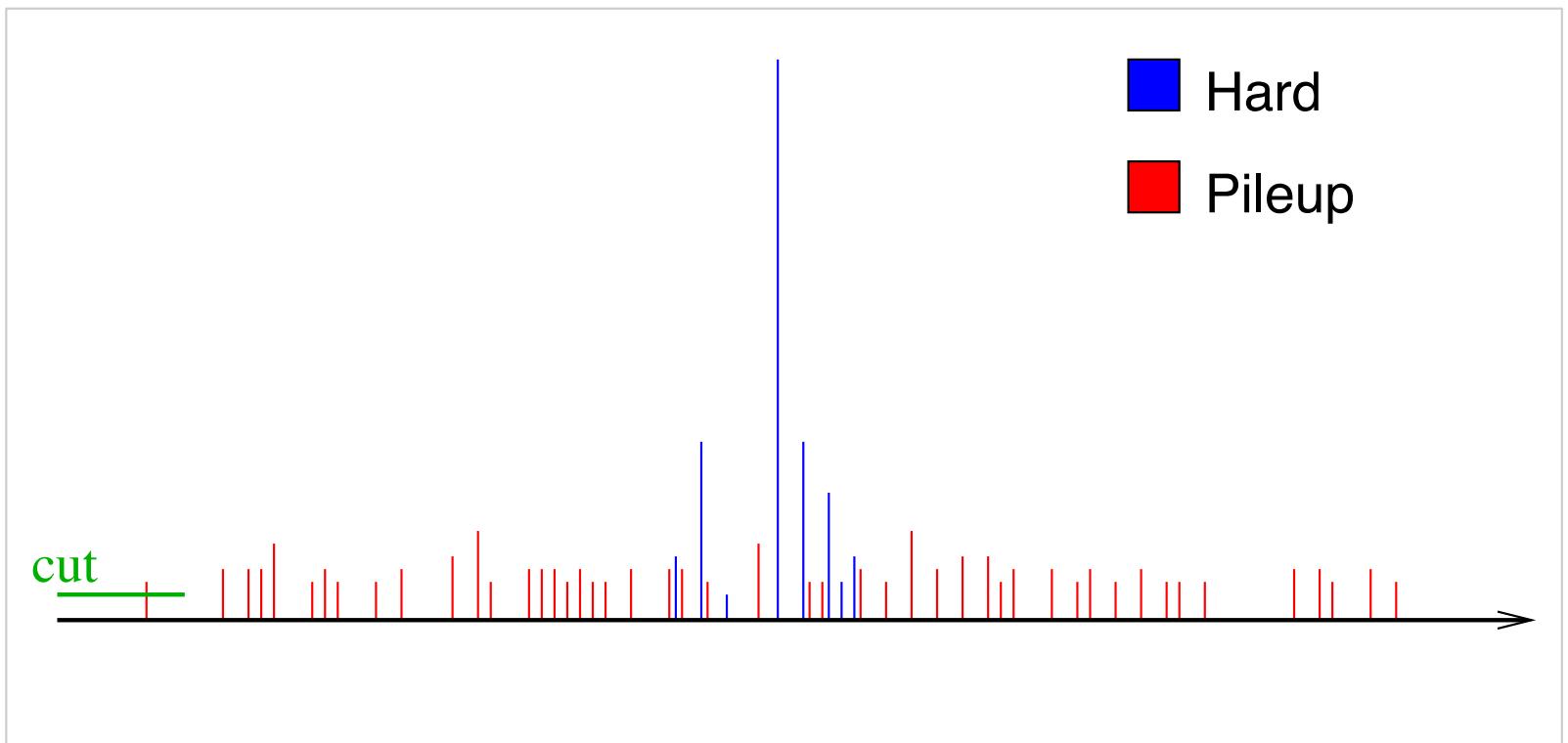
Recover local unbiasedness by rebalancing negative energy flow
into neighbouring areas

Act directly on particles rather than subjets



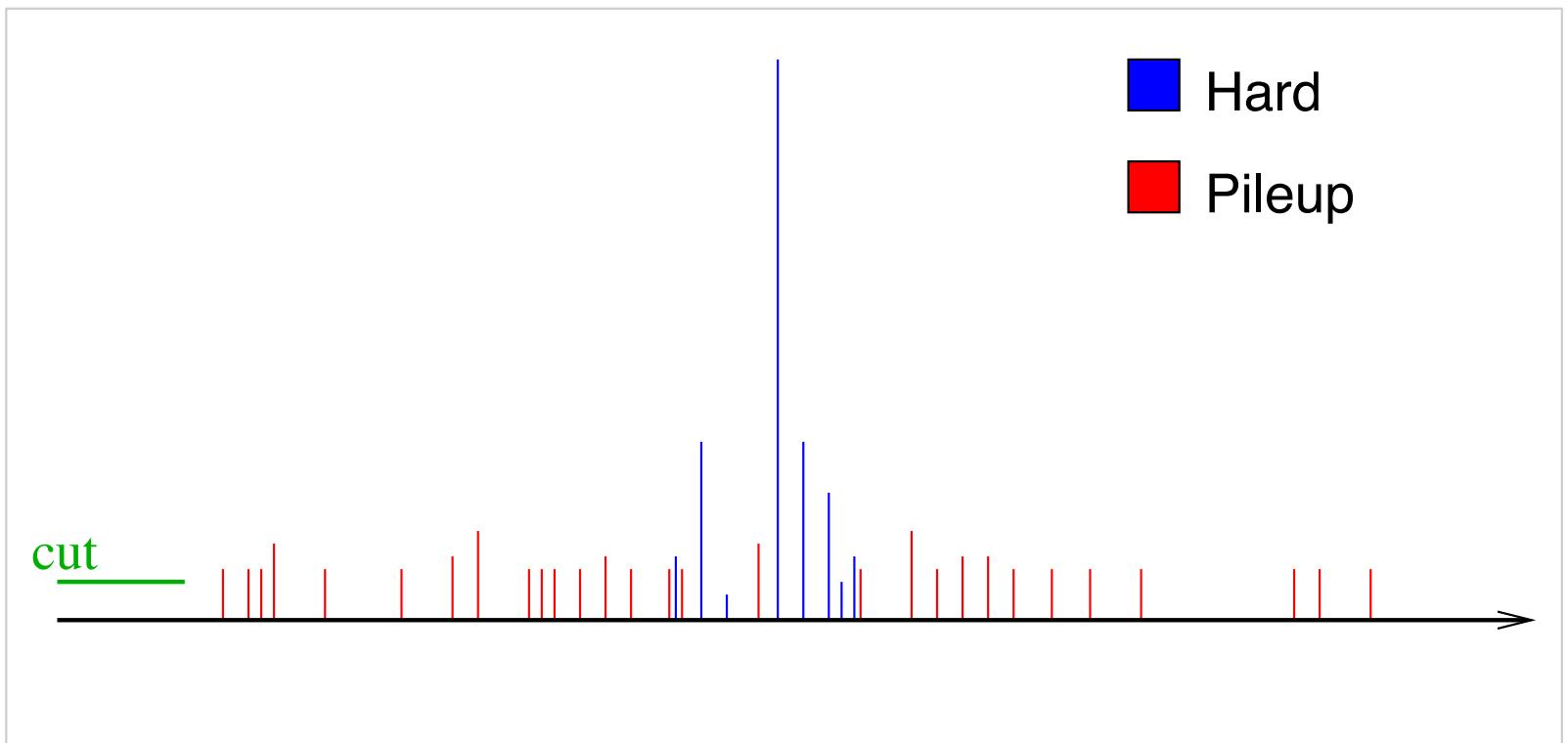
Progressively remove softest particles from event, **until $\rho = 0$**

Soft Killer



G. Soyez

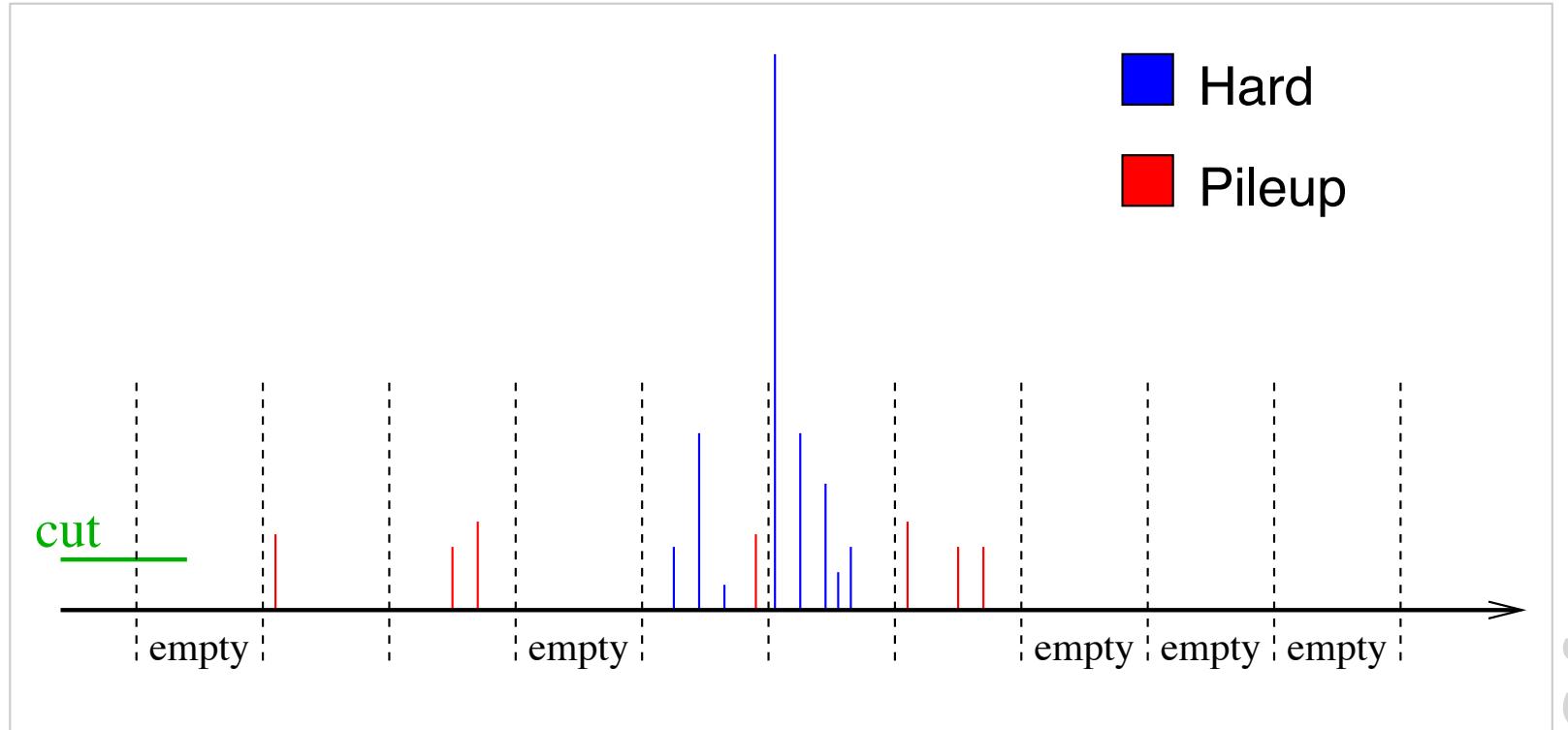
Soft Killer



G. Soyez

Soft Killer

$$p_t^{\text{cut}} = \underset{i \in \text{patches}}{\text{median}} \{p_{ti}^{\max}\}$$



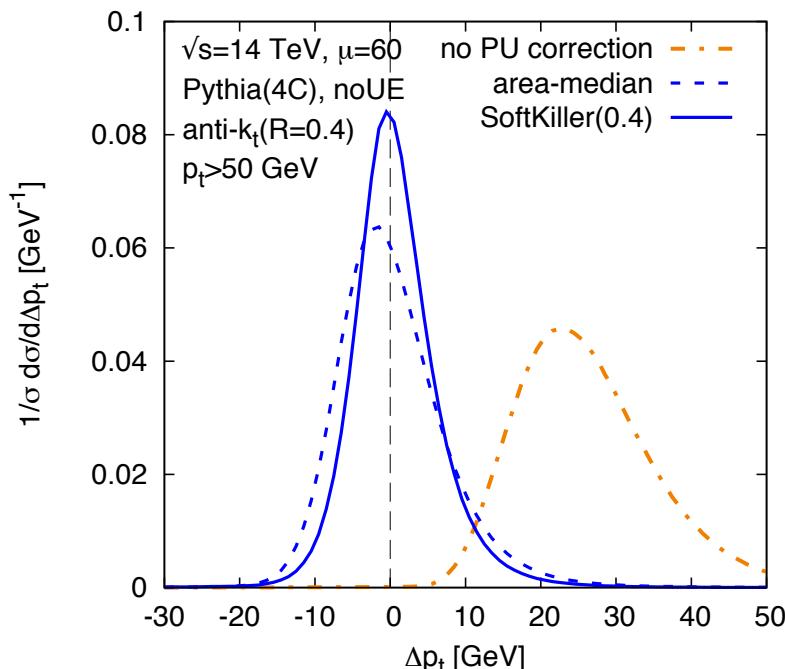
G. Soyez

Half of the event is empty $\Rightarrow \rho = 0$ (because it's the median)

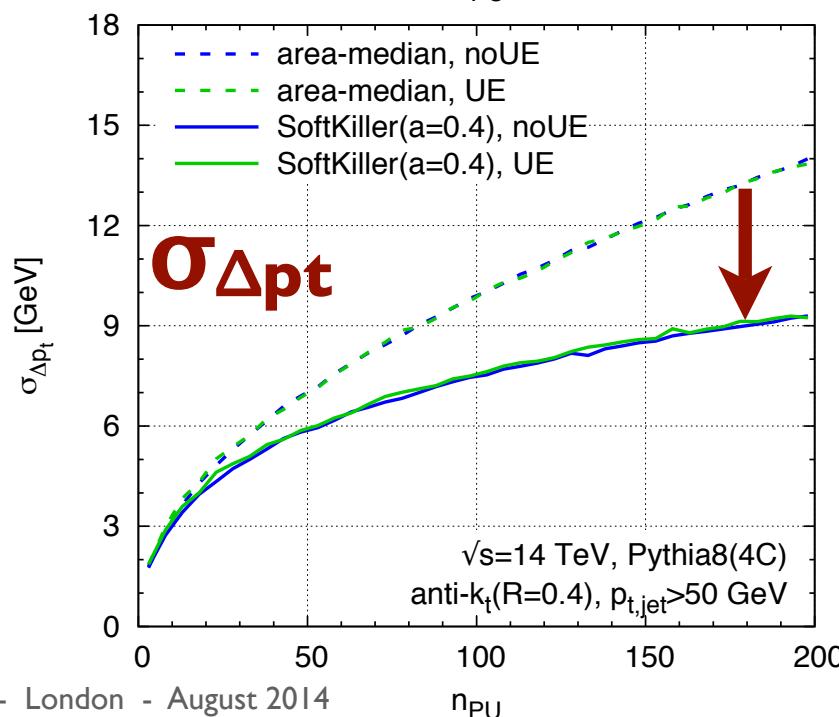
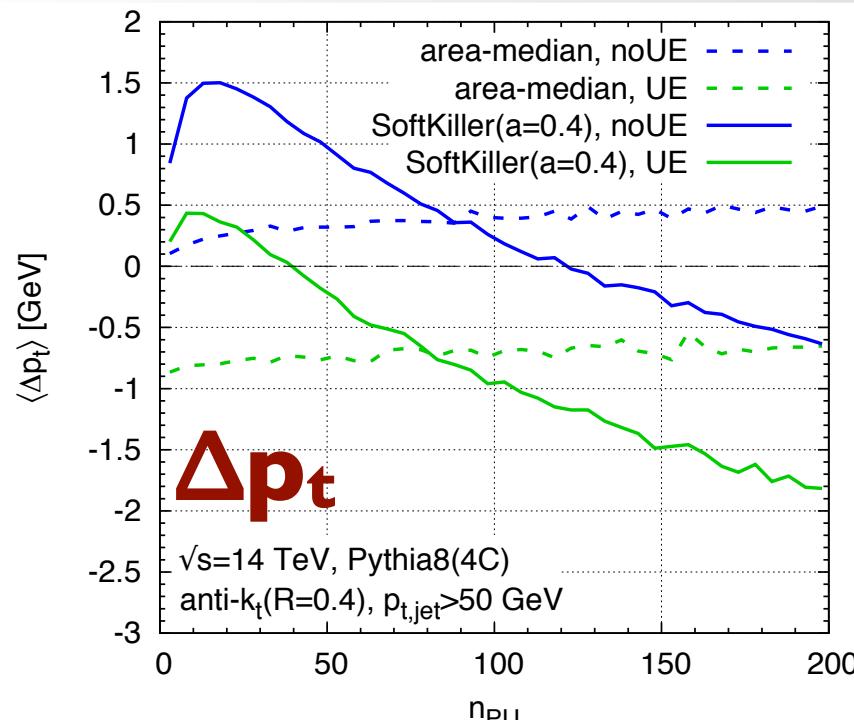
NB. SK needs tuning of the size of the patches used to calculate ρ .
0.4 was found to be a good choice for R=0.4 jets

SoftKiller performance

60 average PU events



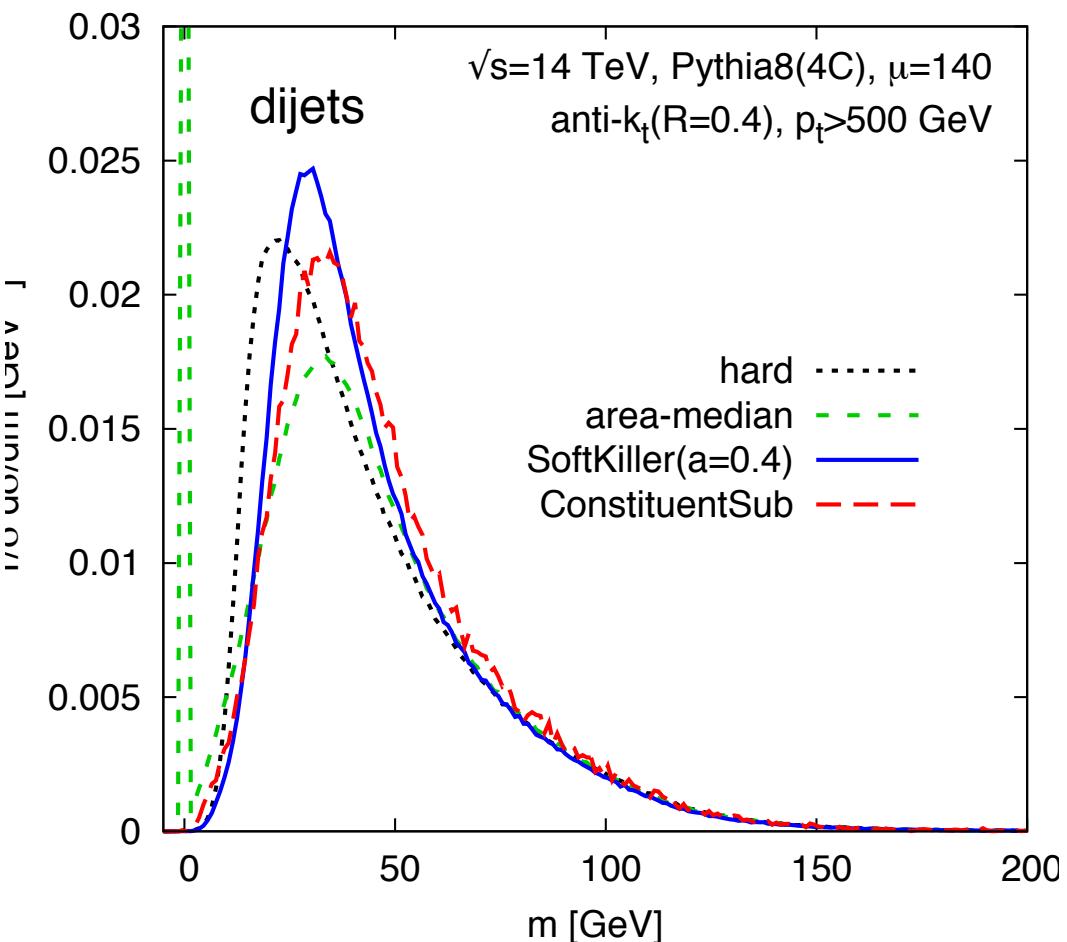
p_t shift, Δp_t



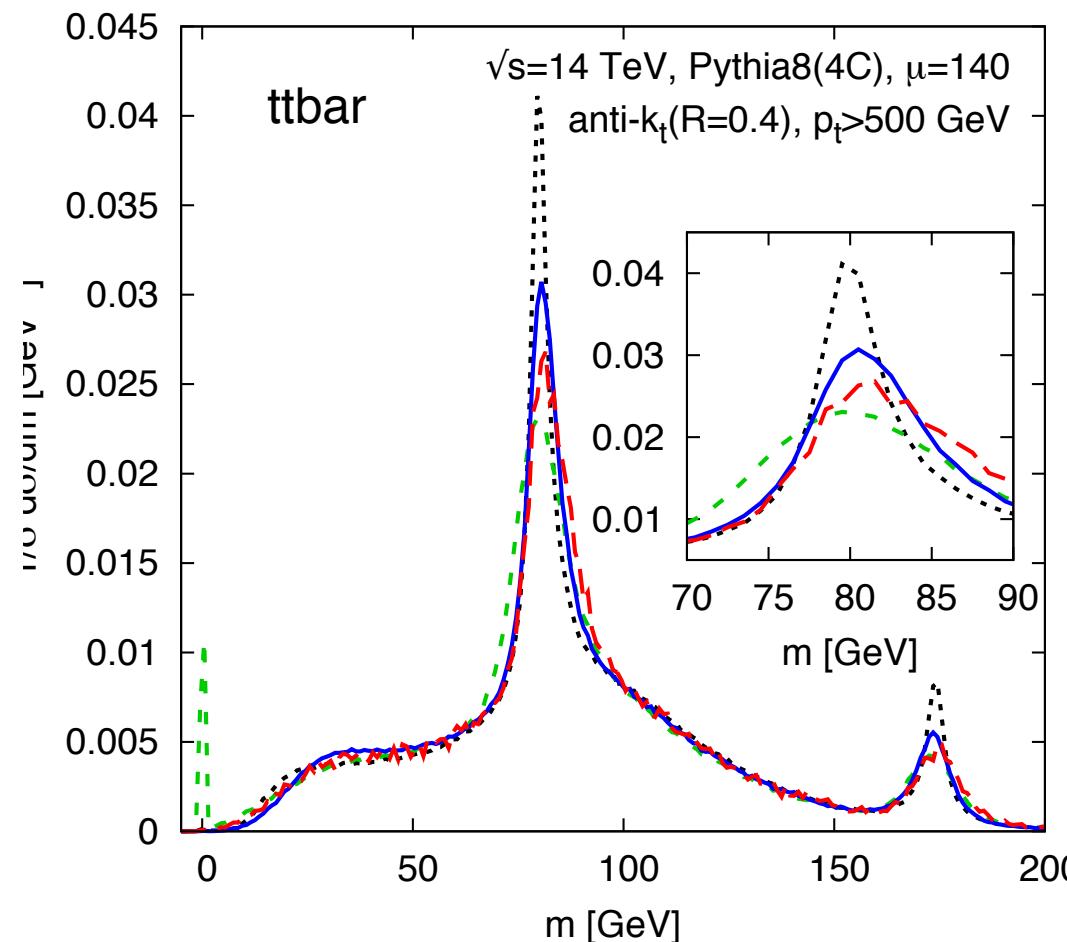
area-median
Soft Killer

Soft Killer performance

Jet mass distributions, 140 average PU events, R=0.4



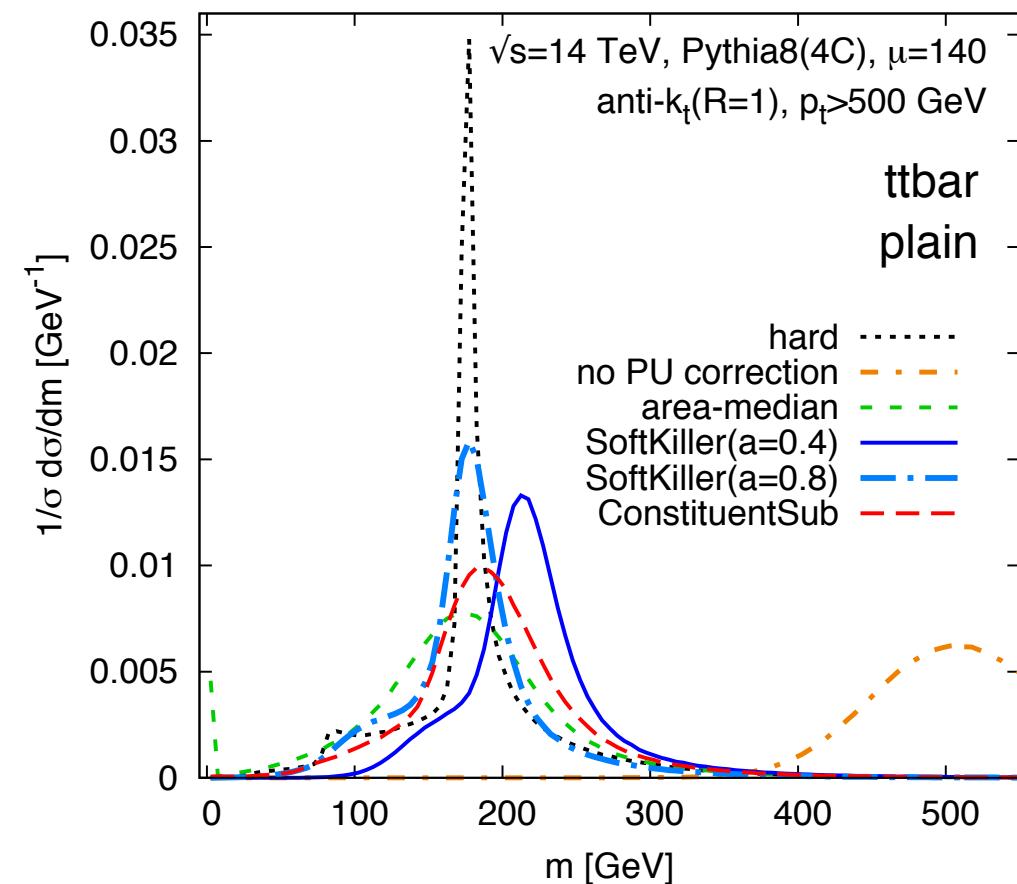
dijet events



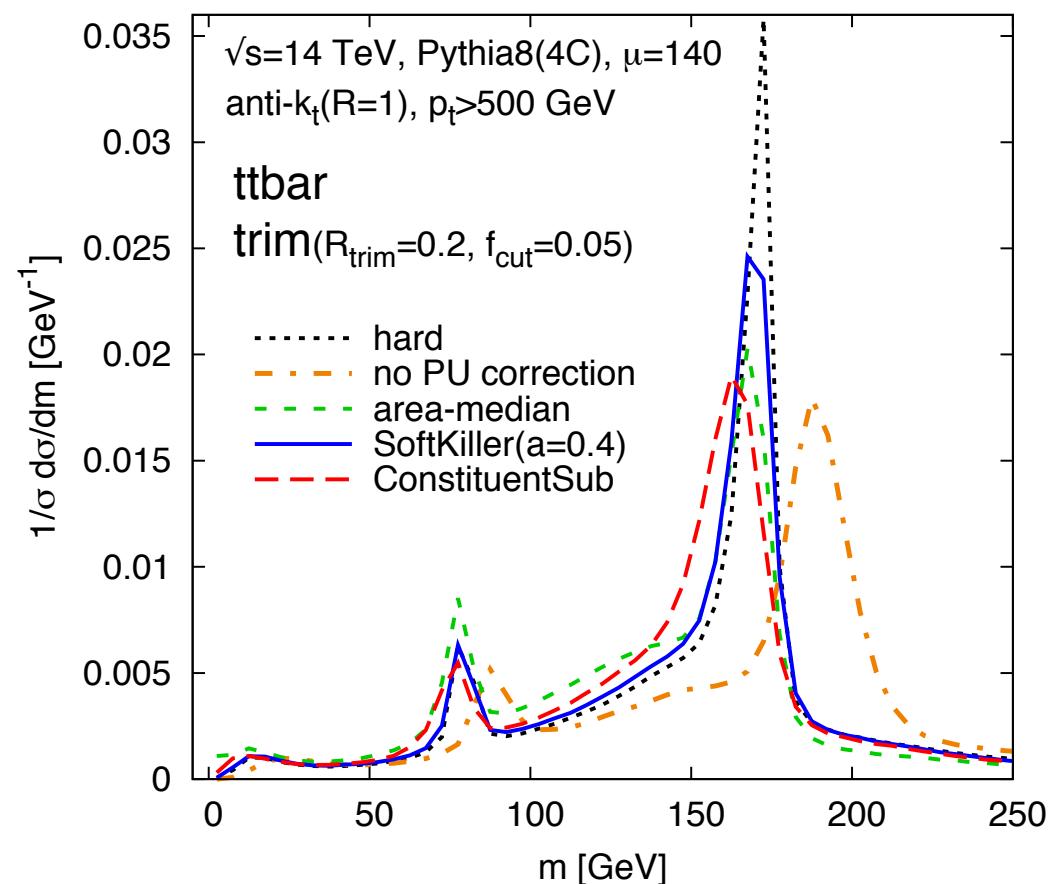
ttbar events

Addition of grooming

SoftKiller with trimming in ttbar events, R=1

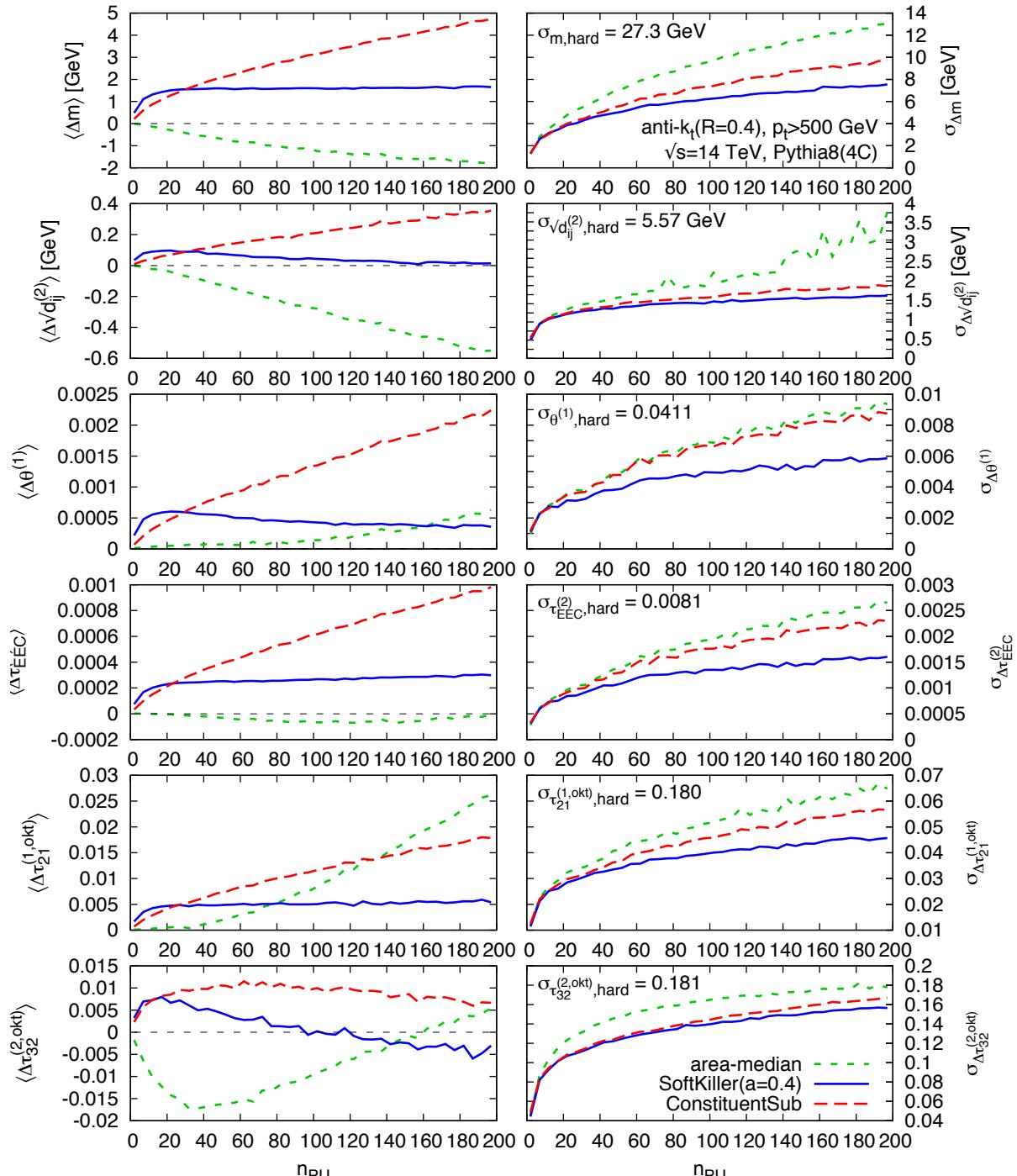


No trimming



Trimming ($R_{\text{sub}}=0.2, f=0.05$)

Soft Killer performance



Many jet shapes:

- jet mass
- kt clustering scale
- jet width (= broadening, = girth)
- energy-energy correlation moment
- T_{21} and τ_{32} N-subjettiness ratios

- Biases under control
- Dispersions smaller than with other methods

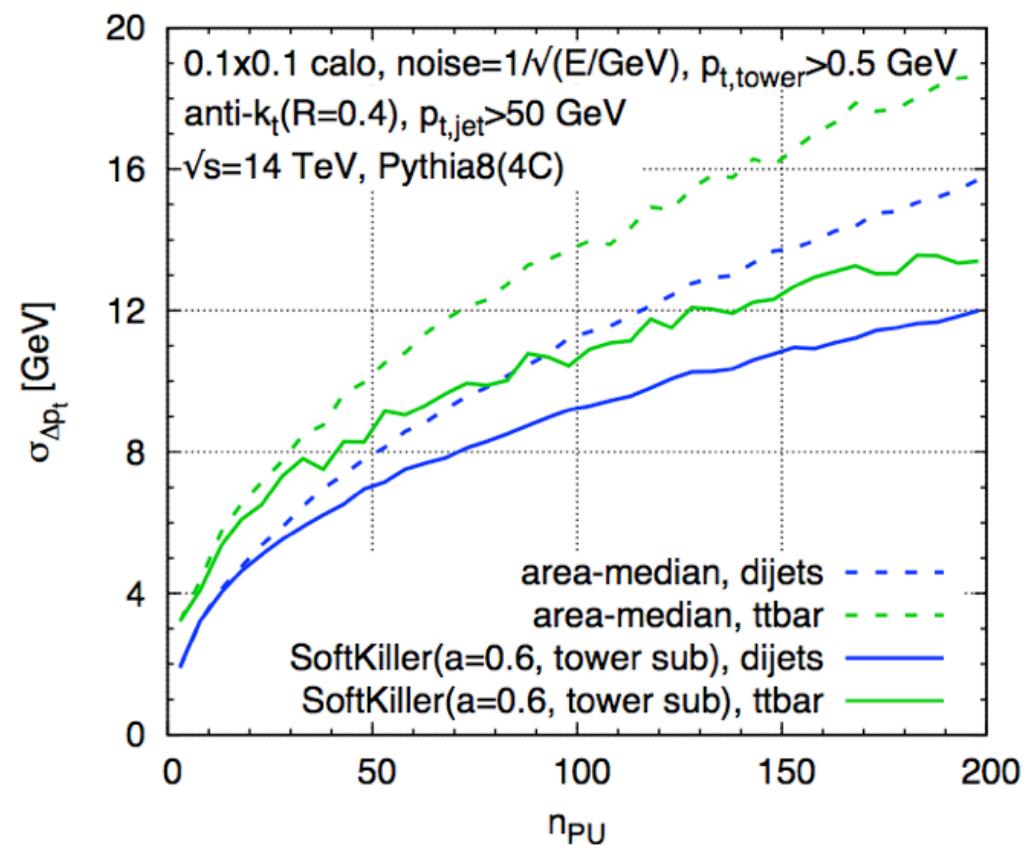
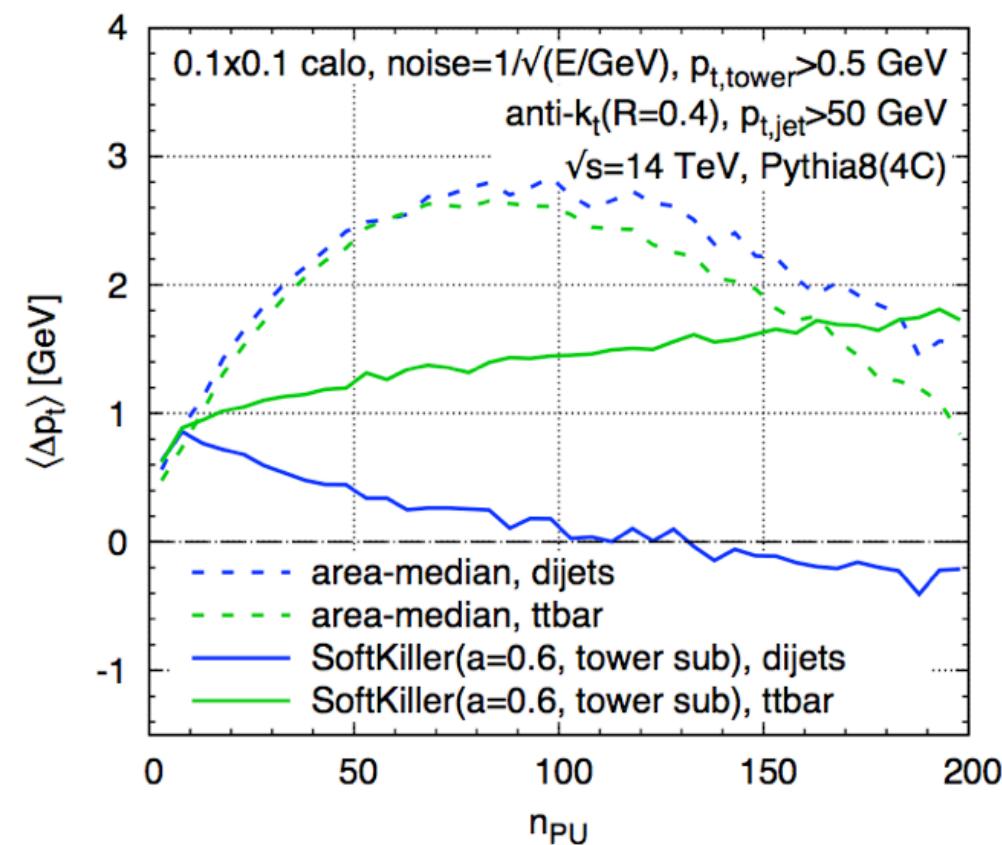
SoftKiller with a calorimeter

Two steps:

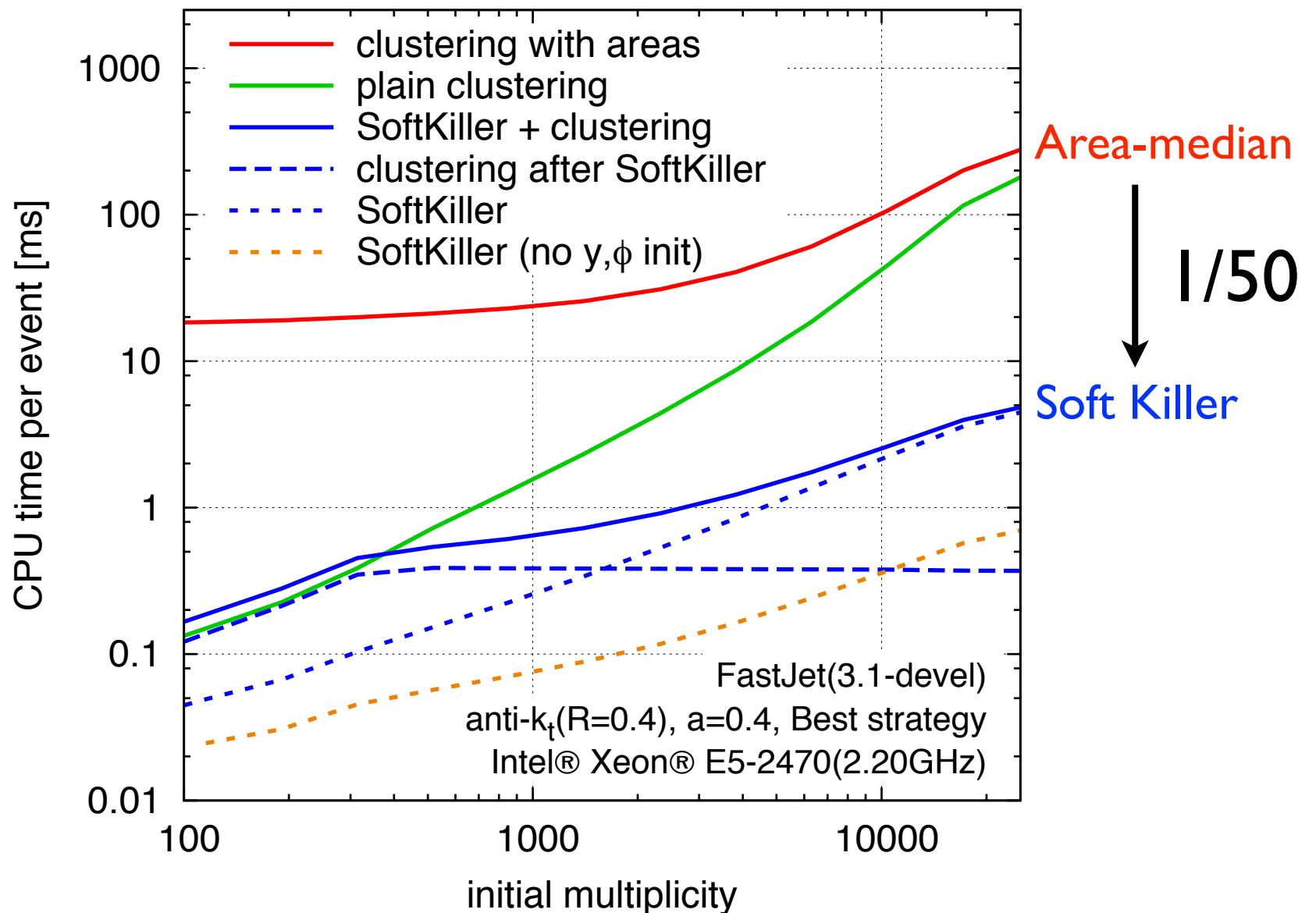
towers area-median subtraction, followed by soft killing

$$p_t^{\text{tower,sub}} = \max(0, p_t^{\text{tower}} - \rho A^{\text{tower}})$$

$$p_t^{\text{cut,sub}} = \underset{i \in \text{patches}}{\text{median}} \left\{ p_{ti}^{\text{tower,sub, max}} \right\}$$

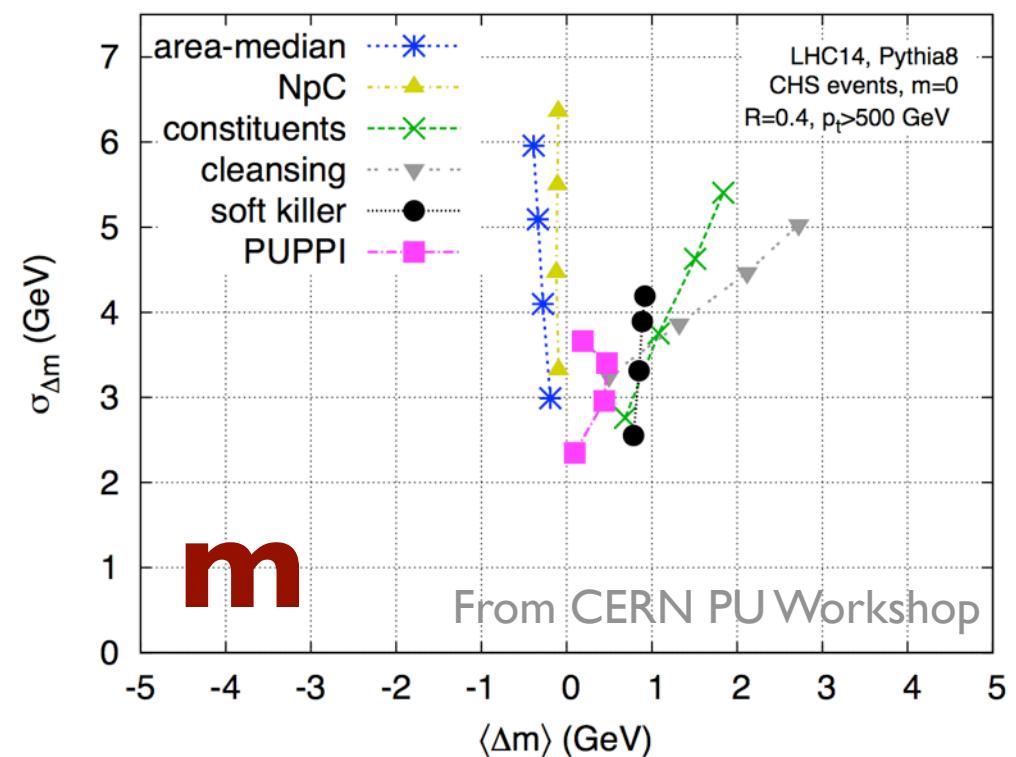
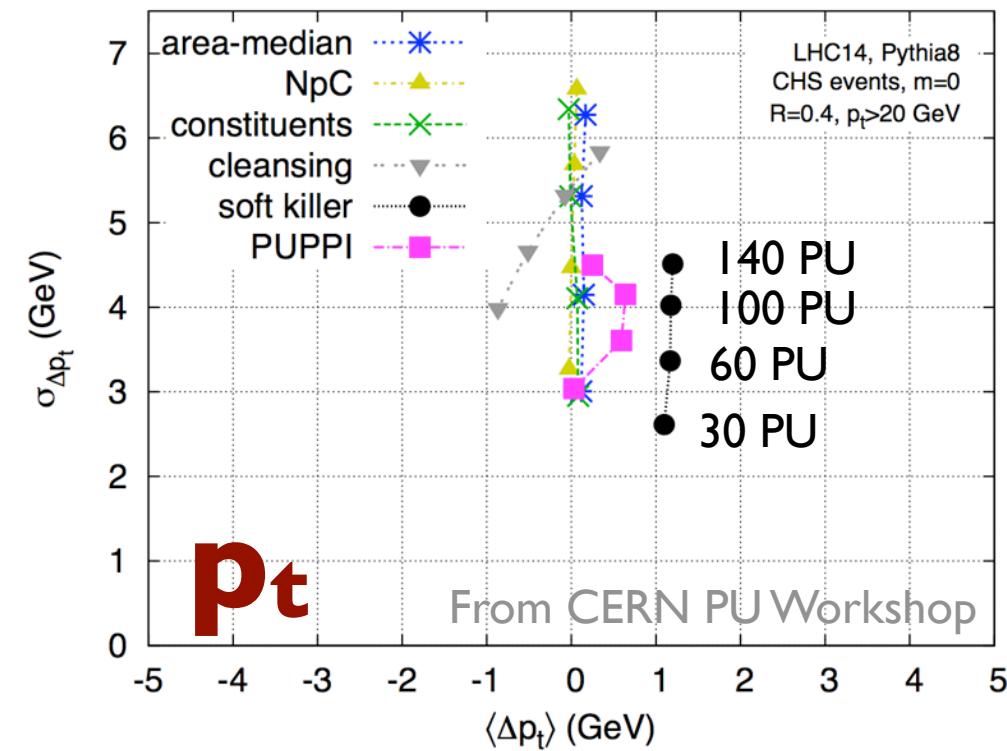


SoftKiller speed



SK is very fast (no clustering is involved).
Almost two orders of magnitude faster than area-median

Comparisons and conclusions



- ▶ Subjet/particle-based background subtraction methods tend to perform better in terms of dispersion than full jet-based ones
 - ▶ can be made reasonably unbiased and robust
 - ▶ can be fast
 - ▶ allow one to calculate any observable
- ▶ Many tools are already public and available in FastJet Contrib