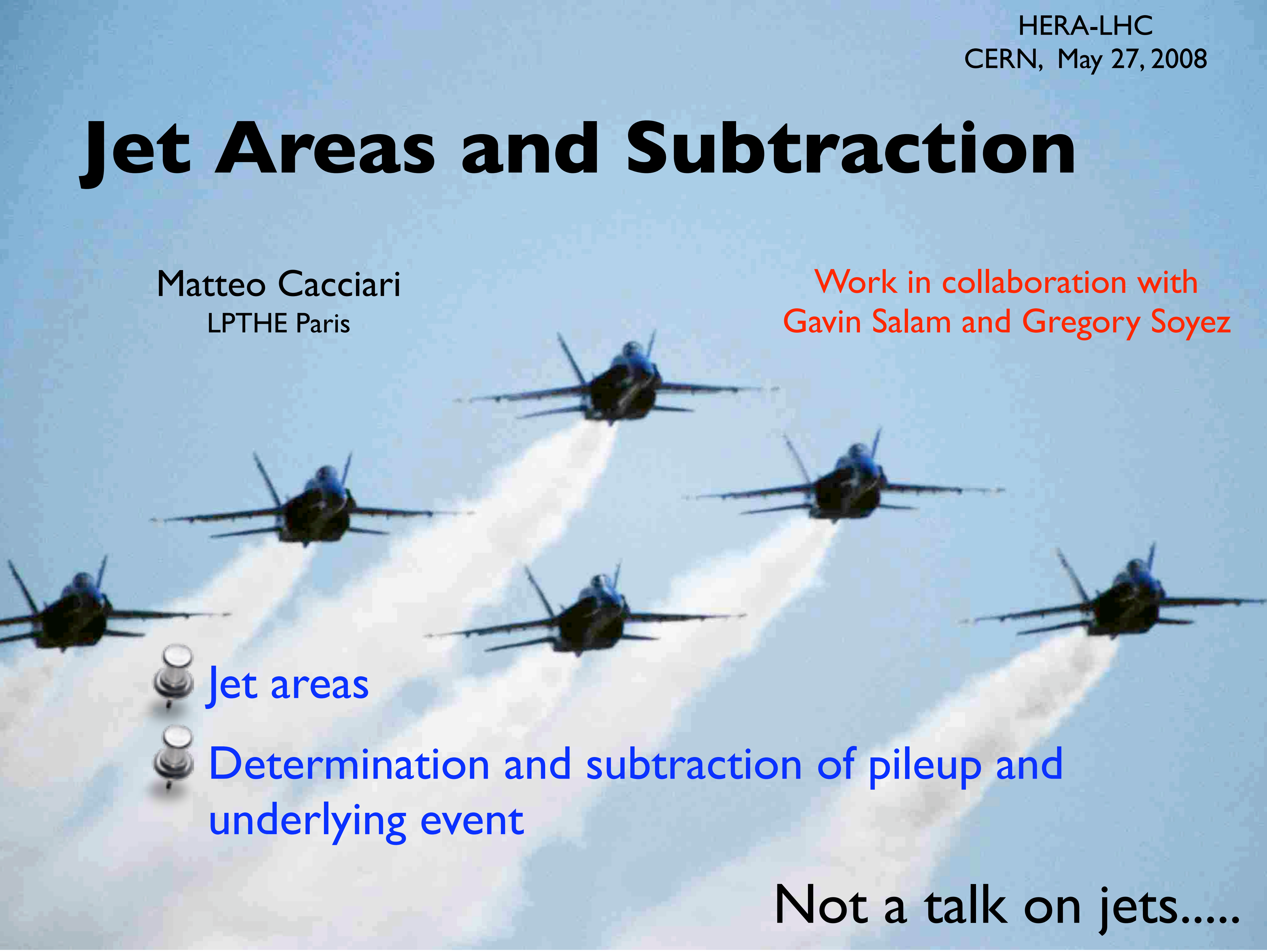
HERA-LHC
CERN, May 27, 2008

Jet Areas and Subtraction

Matteo Cacciari
LPTHE Paris

Work in collaboration with
Gavin Salam and Gregory Soyez

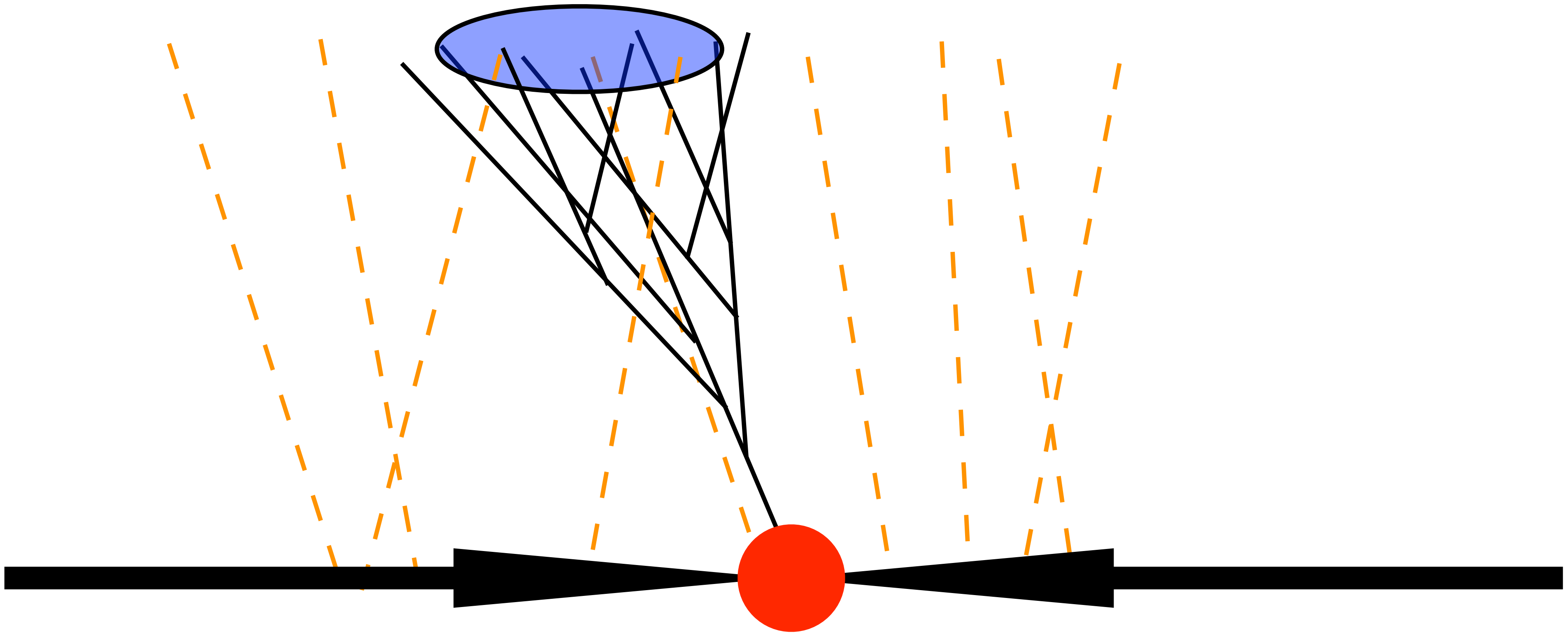


- Jet areas
- Determination and subtraction of pileup and underlying event

Not a talk on jets.....

Making a different use of jets

The physics case

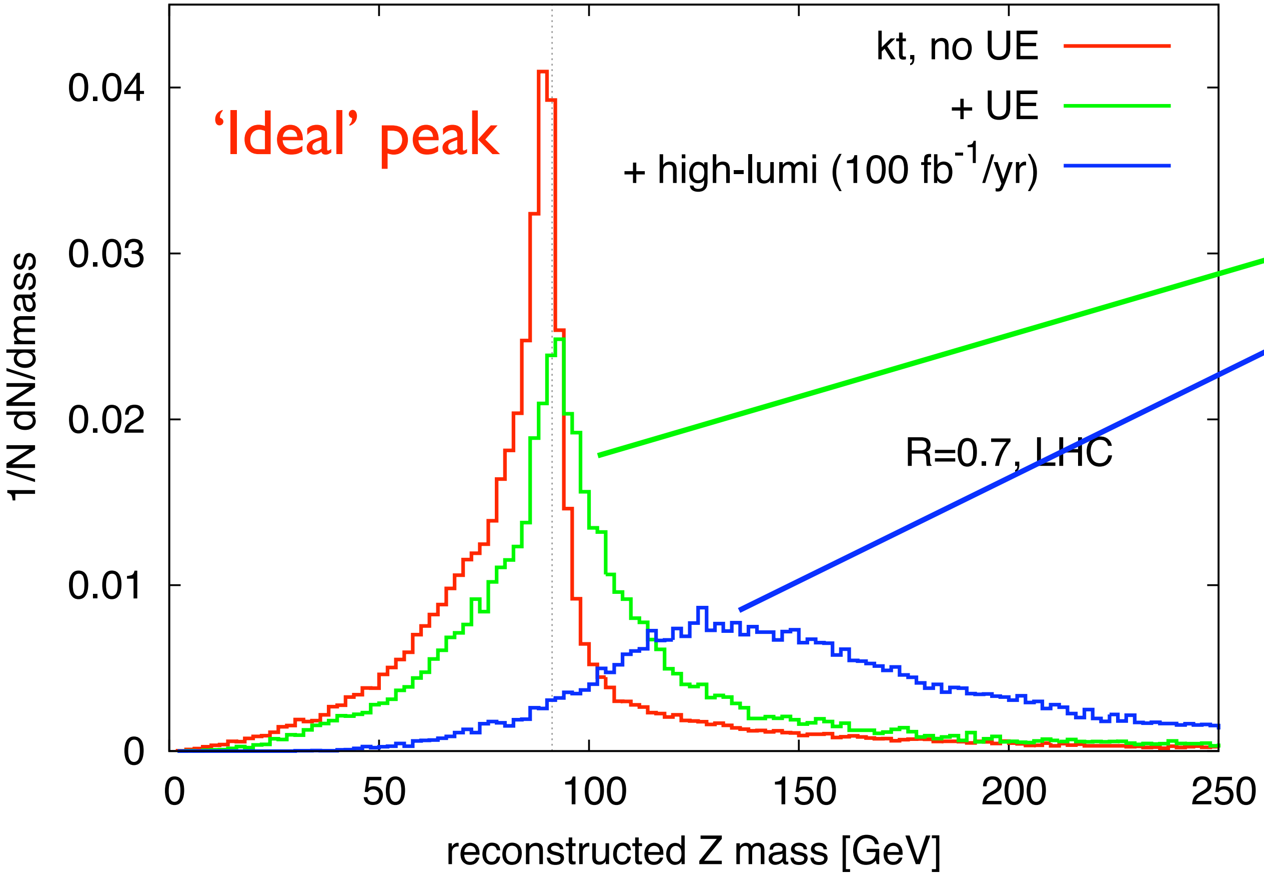


In a realistic set-up underlying event (UE) and pile-up (PU) from multiple collisions produce many soft particles which can ‘contaminate’ the hard jet

p_T (jet) ~ p_T (parton) + Average underlying momentum density × ‘size’ of the jet

The physics case

Challenge at high-energy/high-luminosity machines:
reconstruct objects from jets when a lot of spurious activity is present



Peaks shifted by
UE and pileup.
Resolution degraded

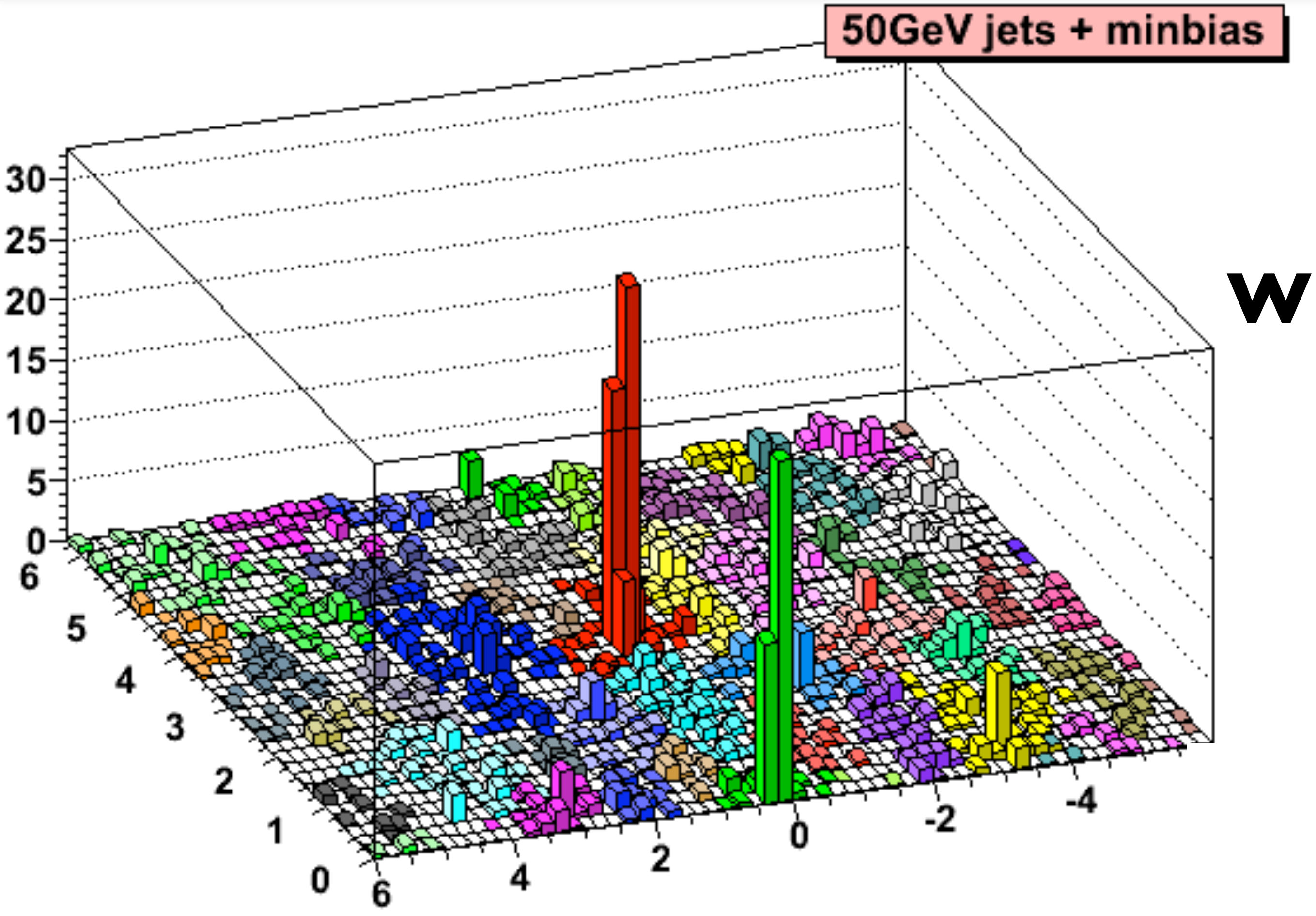
You'd like to be able to subtract the extra stuff from the jets and get back to the correct Z mass

Can we get to know the momentum density of the UE/PU?

Can we subtract it from the jet to find the 'true' momentum?

But...wait...what is the 'size' of a jet??

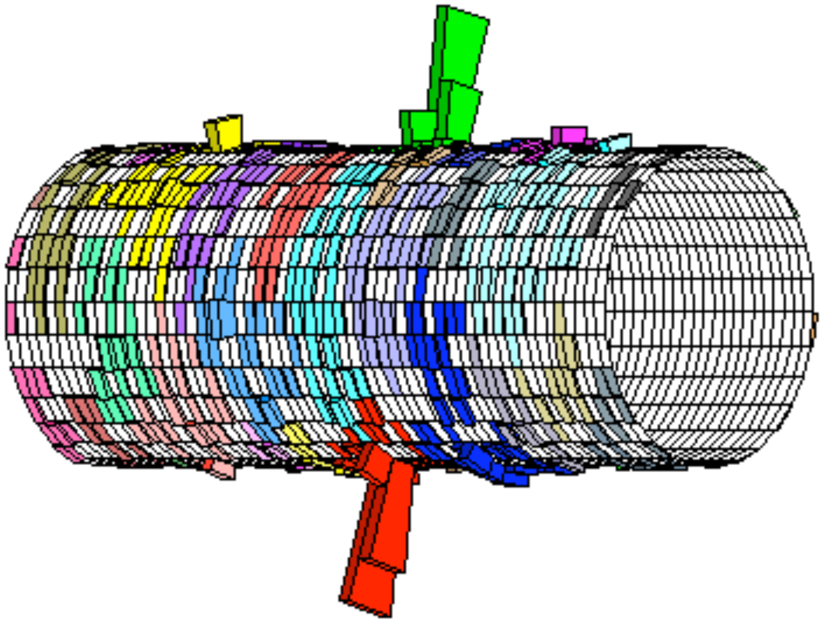
An LHC dijet event



Where does a jet end?

~ 2000 particles

Presence of a jet constituent particle ‘light up’ a cell



Jet areas

Not one, but three **definitions** of a jet's size:

MC, Salam, Soyez, arXiv:0802.1188

- Voronoi area (not discussed here in detail)
- Passive area Mimics effect of **pointlike** radiation (also not discussed here in detail)
- Active area Mimics effect of **diffuse** radiation

(The three areas coincide in the high particle density limit)

[Showing here some theory, but all areas are available natively, for all ICS algorithms and with a user-friendly interface, from **FastJet**, `www.lpthe.jussieu.fr/~salam/fastjet`]

Jet areas

MC, Salam, Soyez, arXiv:0802.1188

Active Area

Add **many** ghost particles in random configurations to the event.
Cluster many times.
Count how many ghosts on average get clustered into a given jet J.

$$A(J\,|\,\{g_i\}) = \frac{N_g(J)}{\nu_g}$$

Number of ghosts in jet J

Active area of a single ghosts configuration

Ghost density

$$A(J) = \lim_{\nu_g \longrightarrow \infty} \langle A(J\,|\,\{g_i\}) \rangle_g$$

Active area

Jet areas calculation

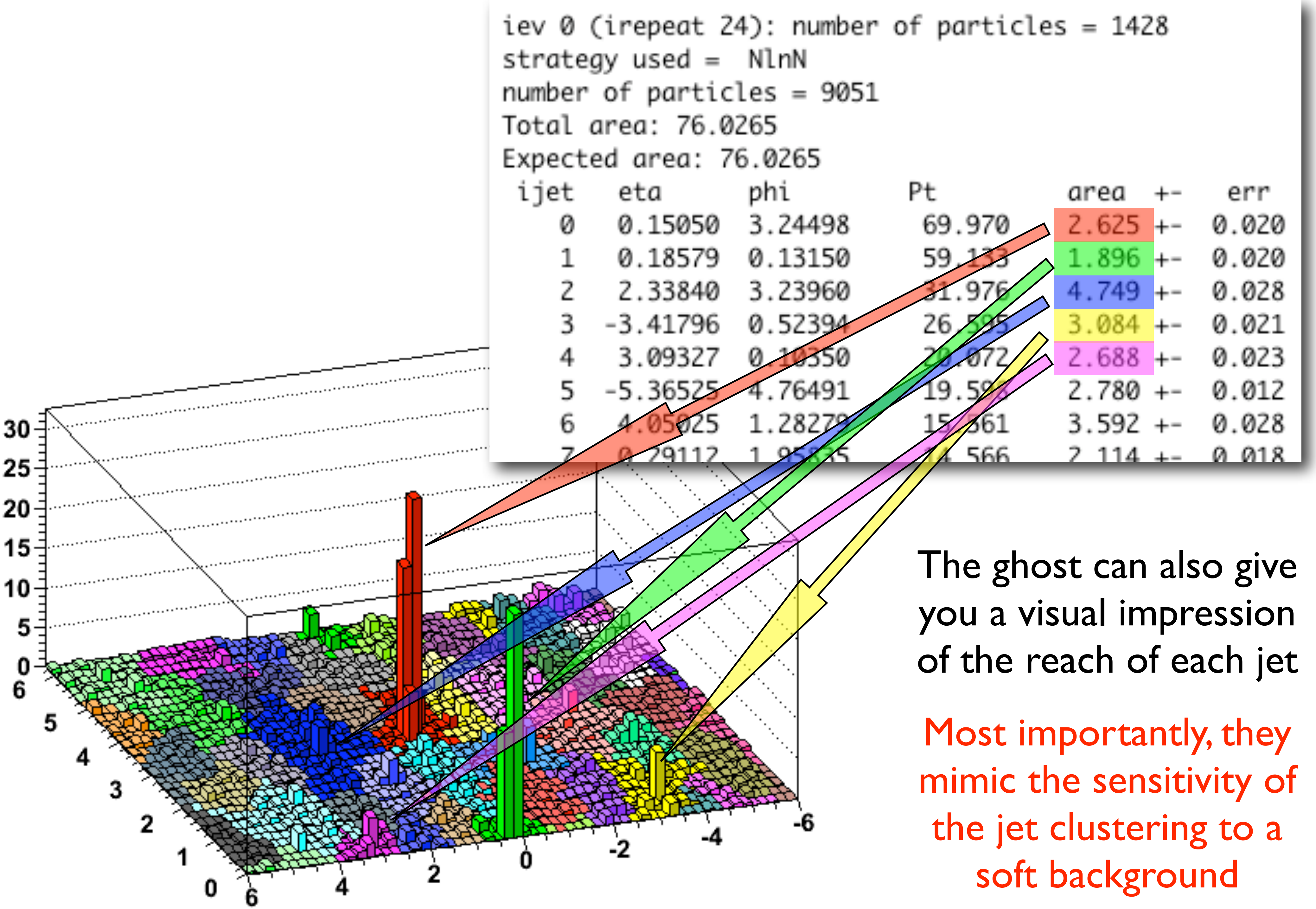
Tools needed to implement it:

1. An **infrared safe jet algorithm** (the ghosts should not change the jets)
2. A reasonably **fast implementation** (we are adding thousands of ghosts)

Both are available

In both cases, determine the area **during** the clustering procedure, **not after it**

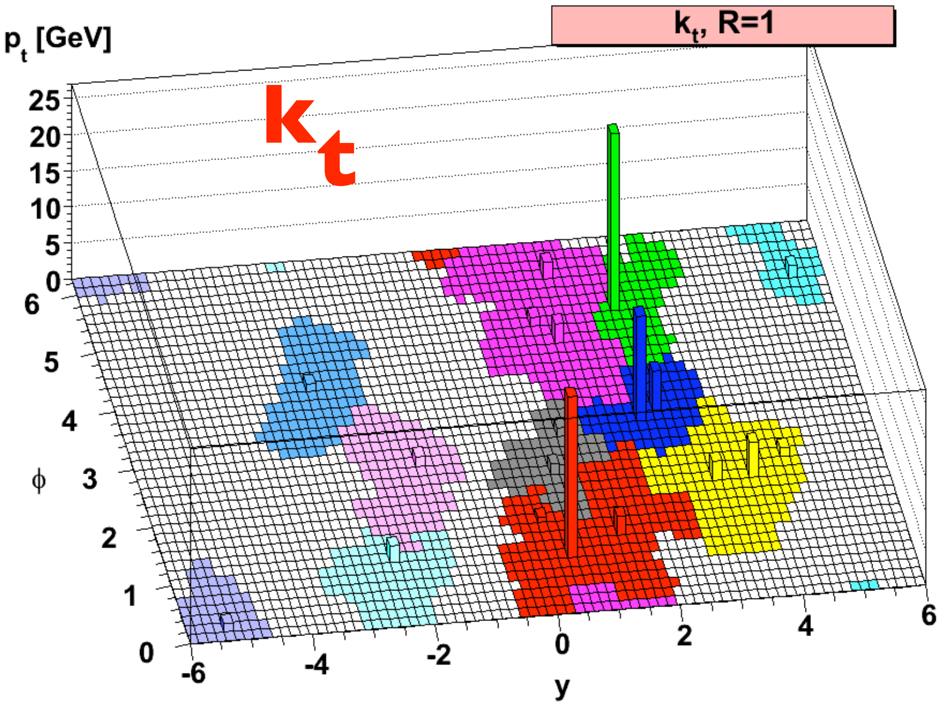
Jet active areas

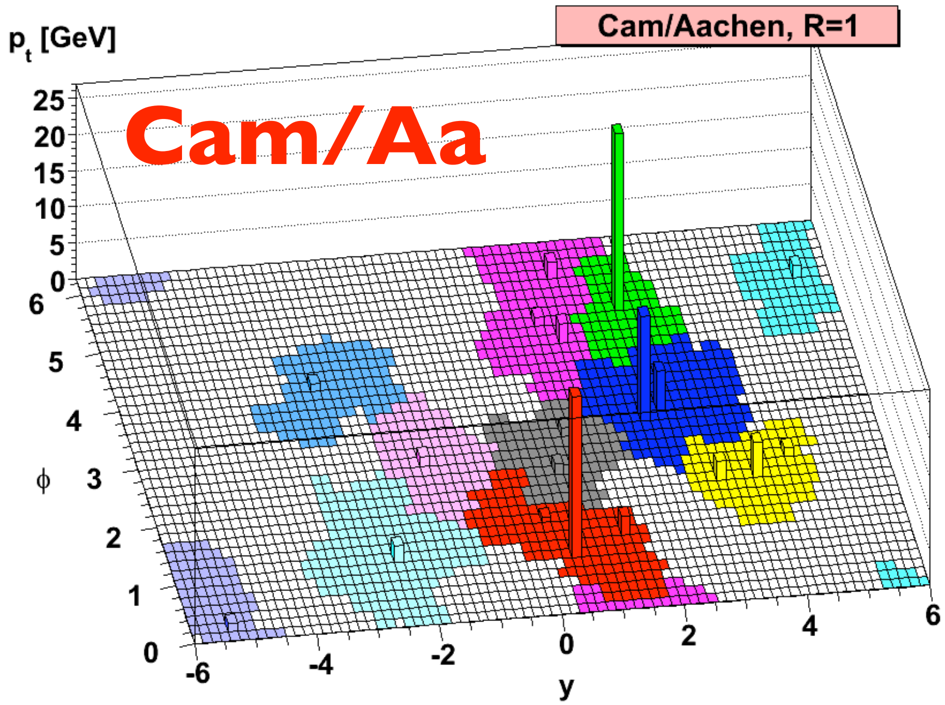


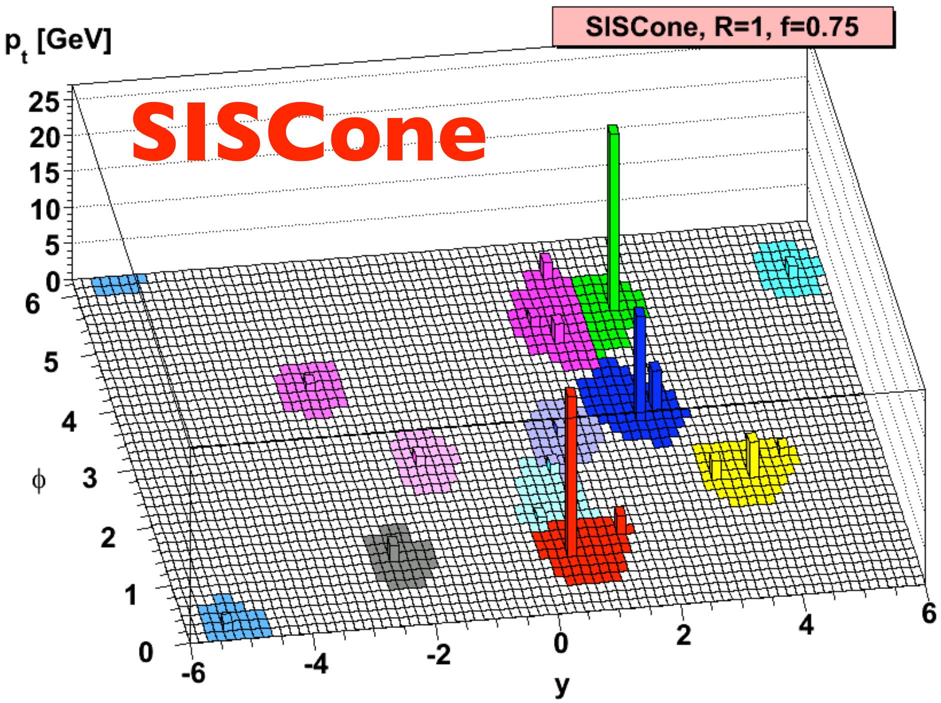
```
iev 0 (irepeat 24): number of particles = 1428
strategy used =  NlnN
number of particles = 9051
Total area: 76.0265
Expected area: 76.0265
 ijet     eta       phi         Pt          area  +-   err
    0   0.15050  3.24498     69.970       2.625 +-  0.020
    1   0.18579  0.13150     59.133       1.896 +-  0.020
    2   2.33840  3.23960     31.976       4.749 +-  0.028
    3  -3.41796  0.52394     26.595       3.084 +-  0.021
    4   3.09327  0.10350     20.072       2.688 +-  0.023
    5  -5.36525  4.76491     19.593       2.780 +-  0.012
    6   4.05025  1.28279     15.561       3.592 +-  0.028
    7   0.29112  1.95835     14.566       2.114 +-  0.018
```

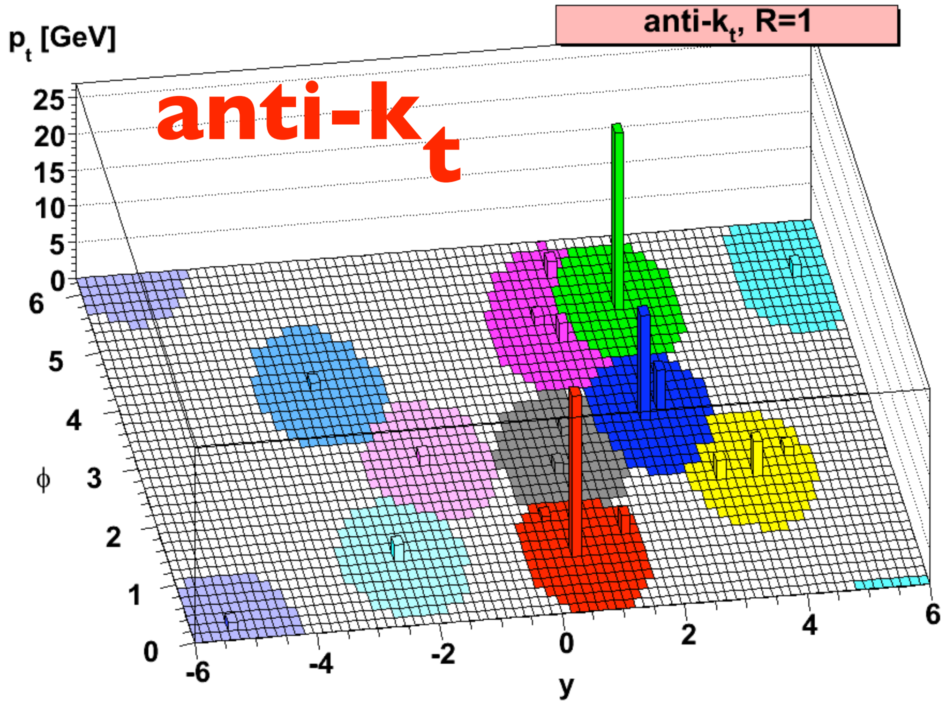
The ghost can also give you a visual impression of the reach of each jet

Most importantly, they mimic the sensitivity of the jet clustering to a soft background


kt, R=1
k_t
p_t [GeV]
25
20
15
10
5
0
6
5
4
ϕ
3
2
1
0
-6
-4
-2
0
2
4
6
y


Cam/Aachen, R=1
Cam/Aa
p_t [GeV]
25
20
15
10
5
0
6
5
4
ϕ
3
2
1
0
-6
-4
-2
0
2
4
6
y


SISCone, R=1, f=0.75
SISCone
p_t [GeV]
25
20
15
10
5
0
6
5
4
ϕ
3
2
1
0
-6
-4
-2
0
2
4
6
y


anti-k_t, R=1
anti-k_t
p_t [GeV]
25
20
15
10
5
0
6
5
4
ϕ
3
2
1
0
-6
-4
-2
0
2
4
6
y

Dispelling the cone-is-a-circle myth

A jet of 'radius' R will surely have area πR^2, right?

Well, it depends.....

Passive areas of a single hard particle are indeed $\boldsymbol{\pi R^2}$

However, **active** areas are not:

- k_t → $0.81\ \pi R^2$
- Cam/Aa → $0.81\ \pi R^2$
- SISCone → $\pi R^2 / 4$
- anti-k_t → πR^2

Recall that 'area' is how much rubbish a jet can pick up.
Its knowledge is essential in order to subtract it from measurements

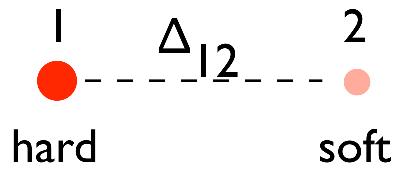
In practice, one calculates numerically with FastJet the area of any given jet

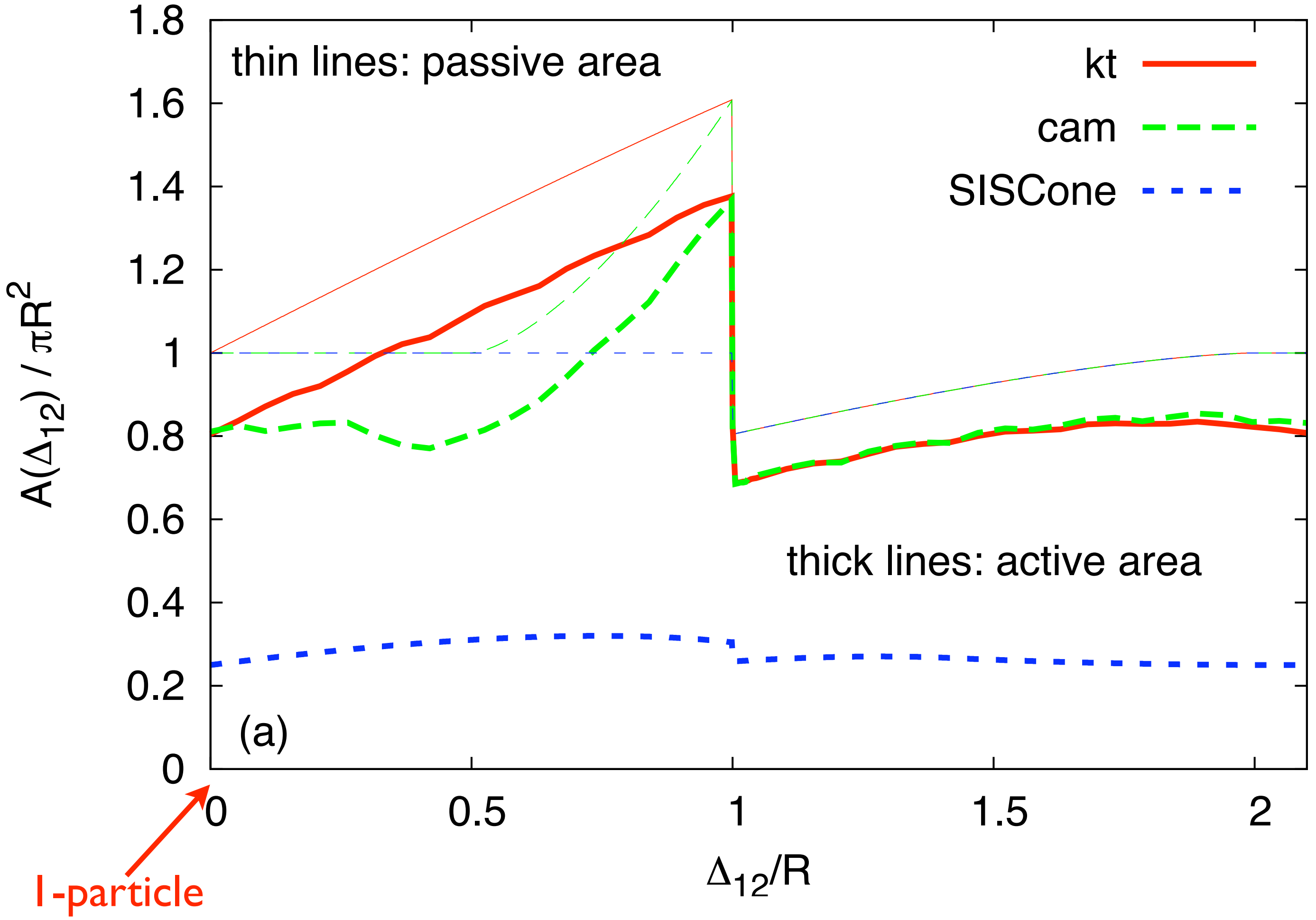
<theory>

Jet areas

Real events have more than a single hard particle.

Add a second (soft) one at a distance Δ_{12}





1-particle limit

The jet area depends on the **distance** between the particles

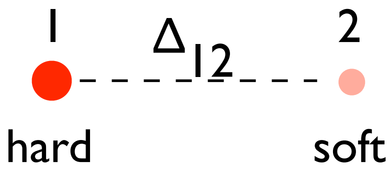
Note very small active area for SISCone!

Passive areas (and SISCone's active area) of jets with two particles (one hard, one soft) can be calculated **analytically**, while the others are obtained numerically

Jet areas

Weigh the probability of emission of the soft particle with the leading QCD matrix element:

$$\langle \Delta area \rangle = \int C_1 \frac{\alpha_s(p_{t2}\Delta_{12})}{\pi} \frac{dp_{t2}}{p_{t2}} \left[\frac{d\Delta_{12}}{\Delta_{12}}\right]_+ \Bigg(\qquad \Bigg)$$



The result is an **anomalous dimension**:

areas change with transverse momentum of the jet in a predictable way:

$$\boxed{\langle \Delta area \rangle = \text{d} \; \frac{C_1}{\pi b_0} \ln \frac{\alpha_s(Q_0)}{\alpha_s(Rp_{t1})}}$$

In a similar way one can also predict the evolution of the dispersion, calculating

$$\langle \Delta \, area^2 \rangle = \text{s}^2 \; \frac{C_1}{\pi b_0} \ln \frac{\alpha_s(Q_0)}{\alpha_s(Rp_{t1})}$$

Passive areas: analytical results

MC, Salam, Soyez, arXiv:0802.1188

d:

$$d_{k_t,R} = \left(\frac{\sqrt{3}}{8} + \frac{\pi}{3} + \xi\right) R^2 \simeq 0.5638\,\pi R^2\,,$$

$$d_{\mathrm{Cam},R} = \left(\frac{\sqrt{3}}{8} + \frac{\pi}{3} - 2\xi\right) R^2 \simeq 0.07918\,\pi R^2\,,$$

$$d_{\mathrm{SISCone},R} = \left(-\frac{\sqrt{3}}{8} + \frac{\pi}{6} - \xi\right) R^2 \simeq -0.06378\,\pi R^2\,,$$

Negative!

s^2:

$$s^2_{k_t,R} = \left(\frac{\sqrt{3}\pi}{4} - \frac{19}{64} - \frac{15\zeta(3)}{8} + 2\pi\xi\right) R^4 \simeq (0.4499\,\pi R^2)^2\,,$$

$$s^2_{\mathrm{Cam},R} = \left(\frac{\sqrt{3}\pi}{6} - \frac{3}{64} - \frac{\pi^2}{9} - \frac{13\zeta(3)}{12} + \frac{4\pi}{3}\xi\right) R^4 \simeq (0.2438\,\pi R^2)^2\,,$$

$$s^2_{\mathrm{SISCone},R} = \left(\frac{\sqrt{3}\pi}{12} - \frac{15}{64} - \frac{\pi^2}{18} - \frac{13\zeta(3)}{24} + \frac{2\pi}{3}\xi\right) R^4 \simeq (0.09142\,\pi R^2)^2\,.$$

with $\xi \equiv \frac{\psi'(1/6) + \psi'(1/3) - \psi'(2/3) - \psi'(5/6)}{48\sqrt{3}} \simeq 0.507471$

Jet areas

	area/πR^2		dispersion		d or D		s or S	
	passive	active	passive	active	passive	active	passive	active
	$a(\mathrm{1PJ})$	$A(\mathrm{1PJ})$	$\sigma(\mathrm{1PJ})$	$\Sigma(\mathrm{1PJ})$	d	D	s	S
k_t	1	0.81	0	0.28	0.56	0.52	0.45	0.41
Cam/Aachen	1	0.81	0	0.26	0.08	0.08	0.24	0.19
SISCone	1	1/4	0	0	−0.06	0.12	0.09	0.07
anti-k_t	1	1	0	0	0	0	0	0

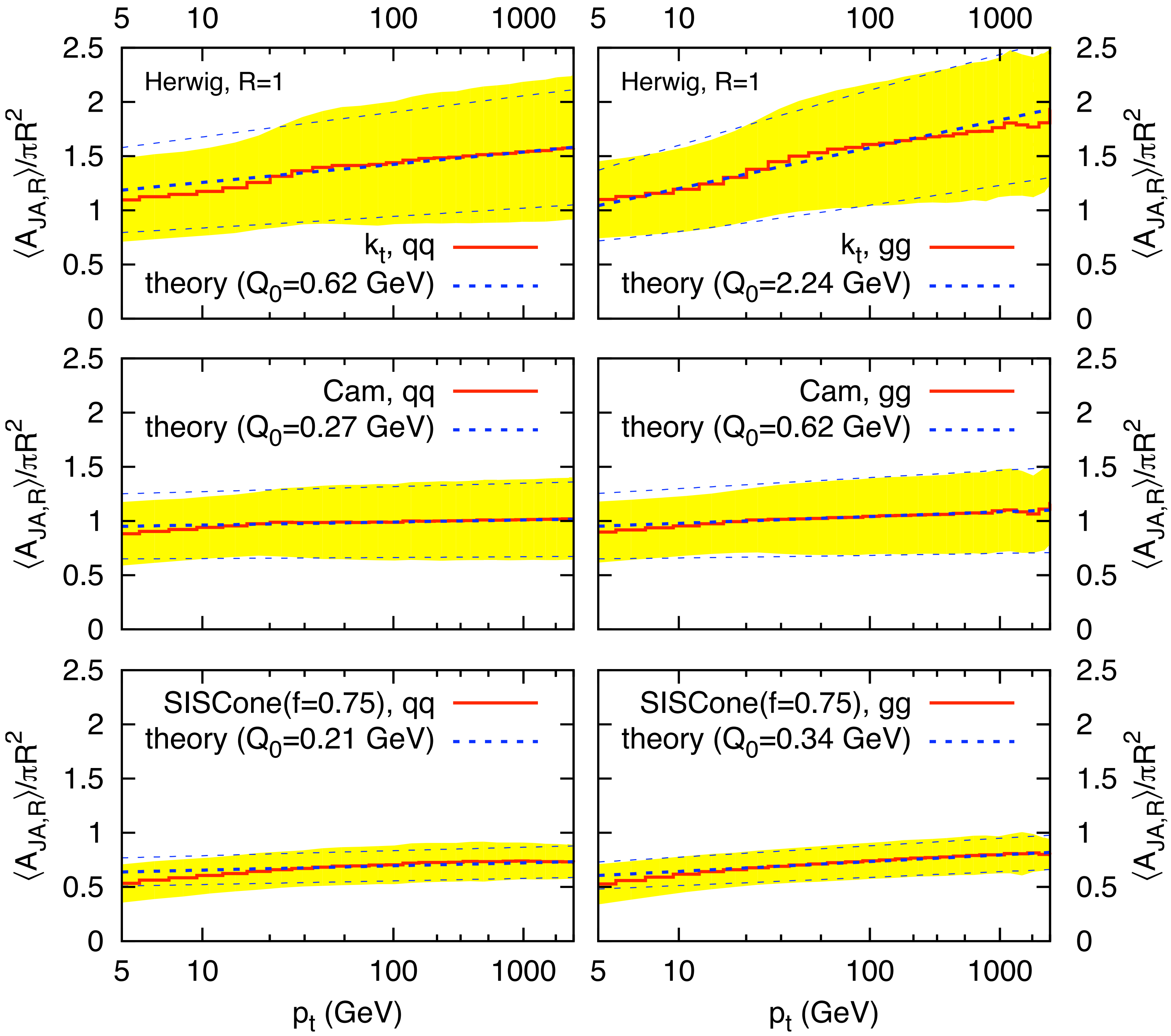
single hard particle

emission of a second perturbative particle (coeff. of anomalous dimension)

Some remarkable features

- SISCone has very small active area
- SISCone's anomalous dimension changes from negative for passive area to positive for active area
- kt has largest anomalous dimension
- anti-kt has constant area (null anomalous dimension): it's a **perfect cone**

Jet area scaling violations at (simulated) LHC



Averages and dispersions evolution from Monte Carlo simulations in good agreement with simple LL calculations

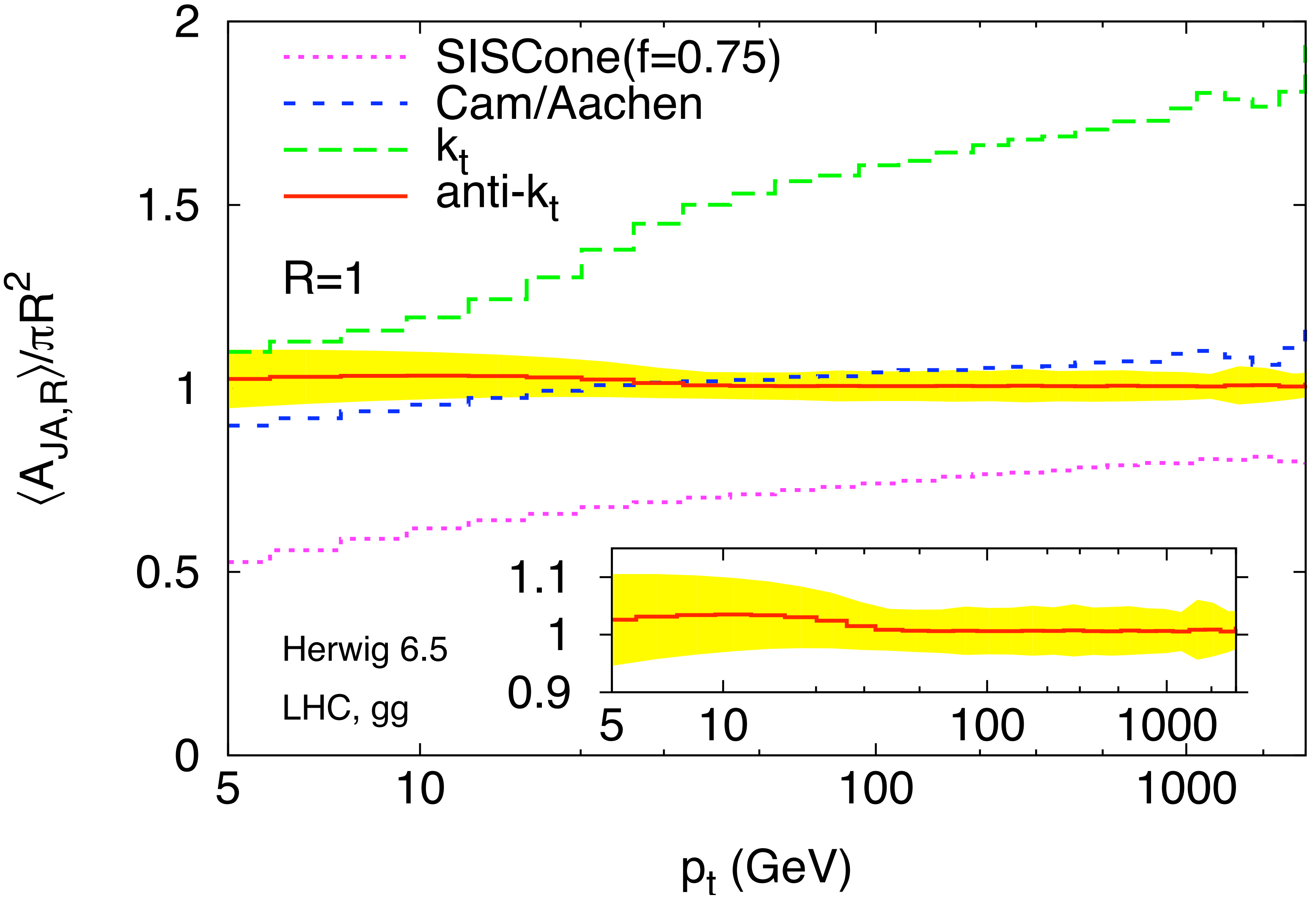
Area scaling violations are a legitimate observable!

(Though it might not be the best place where to measure α_s ....)

Jet area scaling violations at (simulated) LHC

MC, Salam, Soyez, arXiv:0802.1189

Check anti-k_t behaviour: scaling violations indeed absent, as predicted

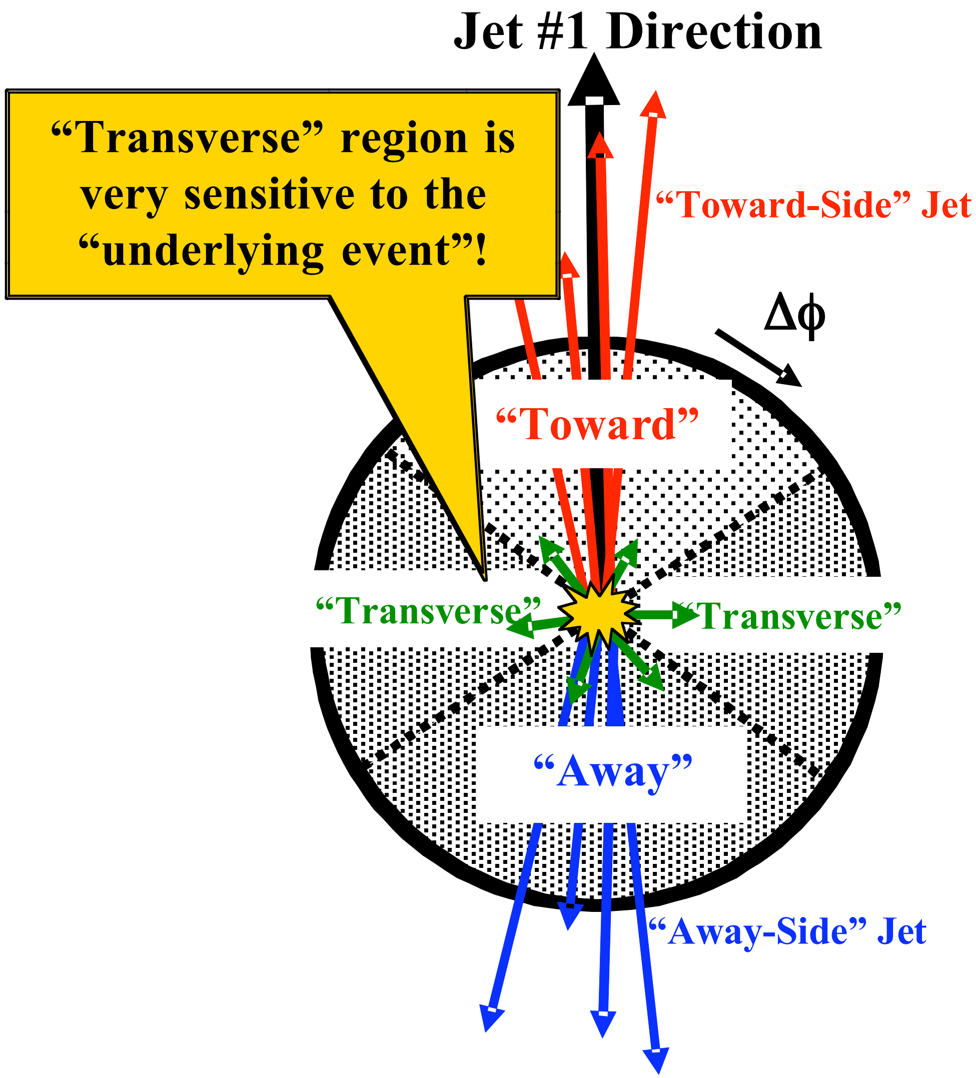


</theory>

Jet areas as a tool:

Underlying event and pileup determination and subtraction

Common approach

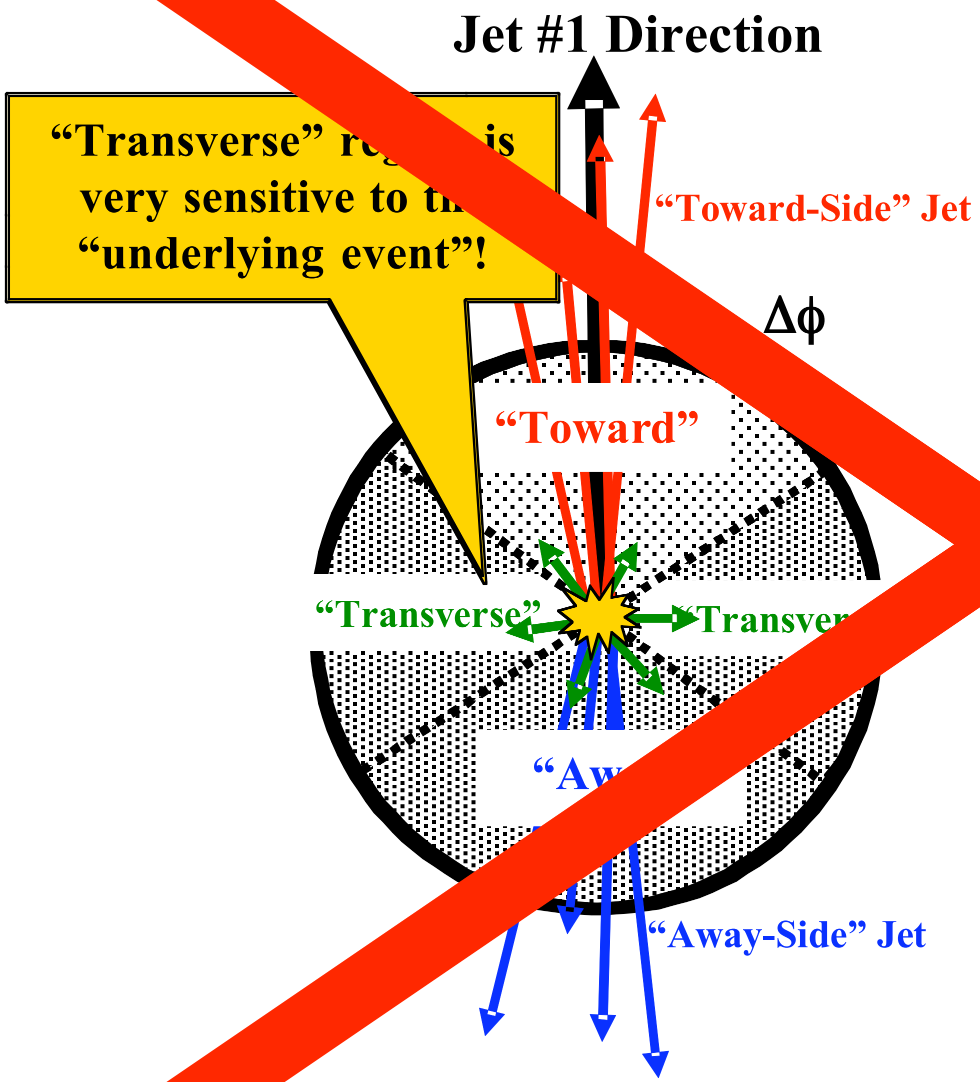


Marchesini-Webber idea:
look at transverse region to
measure underlying event

Topological selection

The jets are classified as belonging to the noise on the ground of their **position**

Common approach



“Transverse” re is very sensitive to th “underlying event”!

Marchesini-We er idea:
look at tran se region to
measu nderlying event

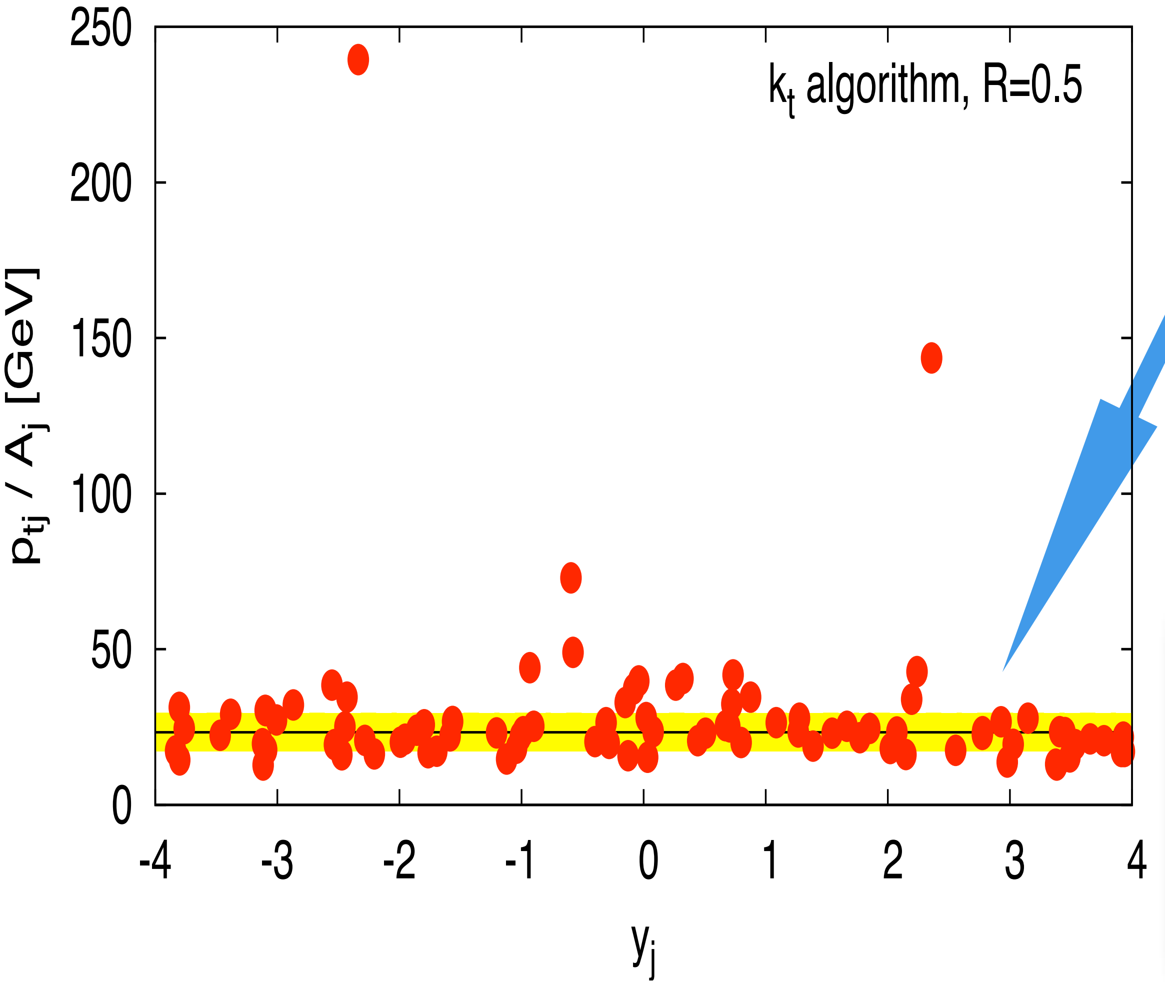
T ological selection

The jet classified as belonging to the no on the ground of their p **ition**

The key observation

LHC: dijet event + high-lumi pilup

p_T/Area is fairly constant, except for the hard jets



The distribution of background jets establishes its own average momentum density **ρ**

(NB. this is true on an event-by-event basis)

NB. This and following examples done with pile-up, but it works similarly with underlying event, to some extent

Dynamical selection

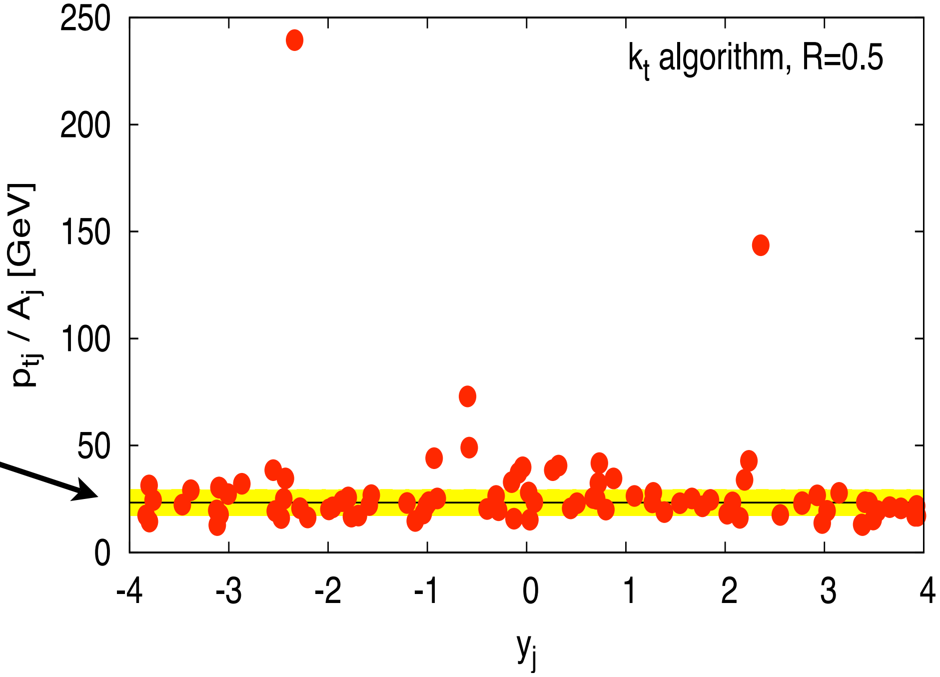
The jets are classified as belonging to the noise on the ground of their **characteristics**

Extraction of average noise momentum density

$$\rho \equiv \text{median}\left[\left\{\frac{p_t^{jet}}{\text{Area}_{jet}}\right\}\right]$$

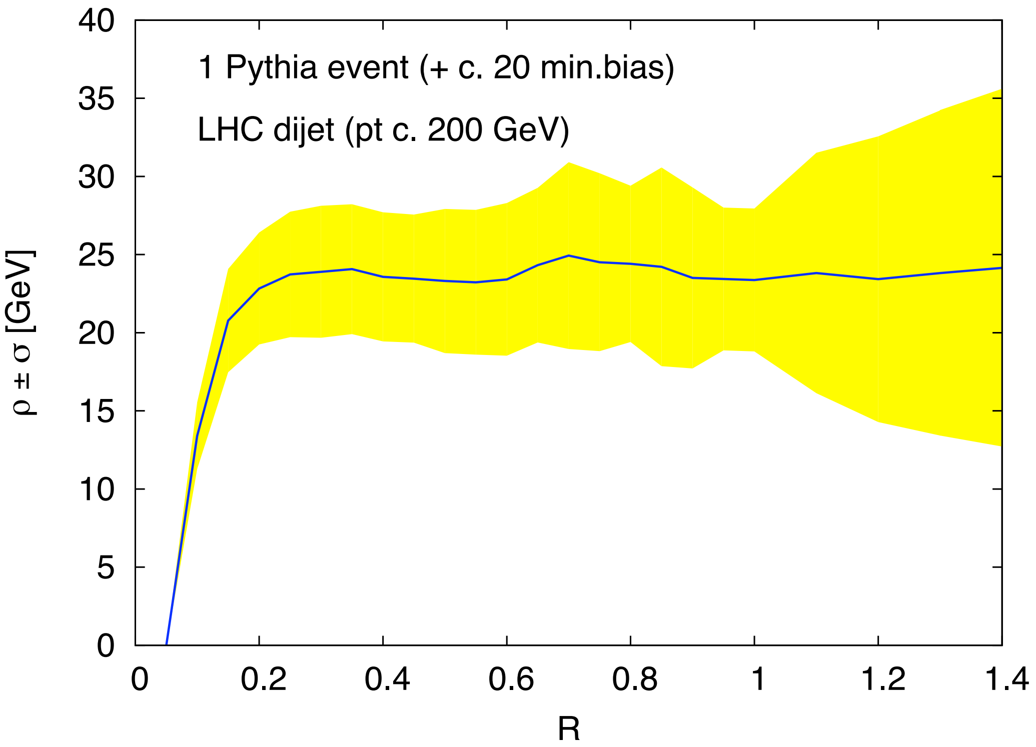
(Taking the median of the distribution is a nice trick to get rid of the possible bias from the few hard jets)

One can also estimate the fluctuations (yellow band)

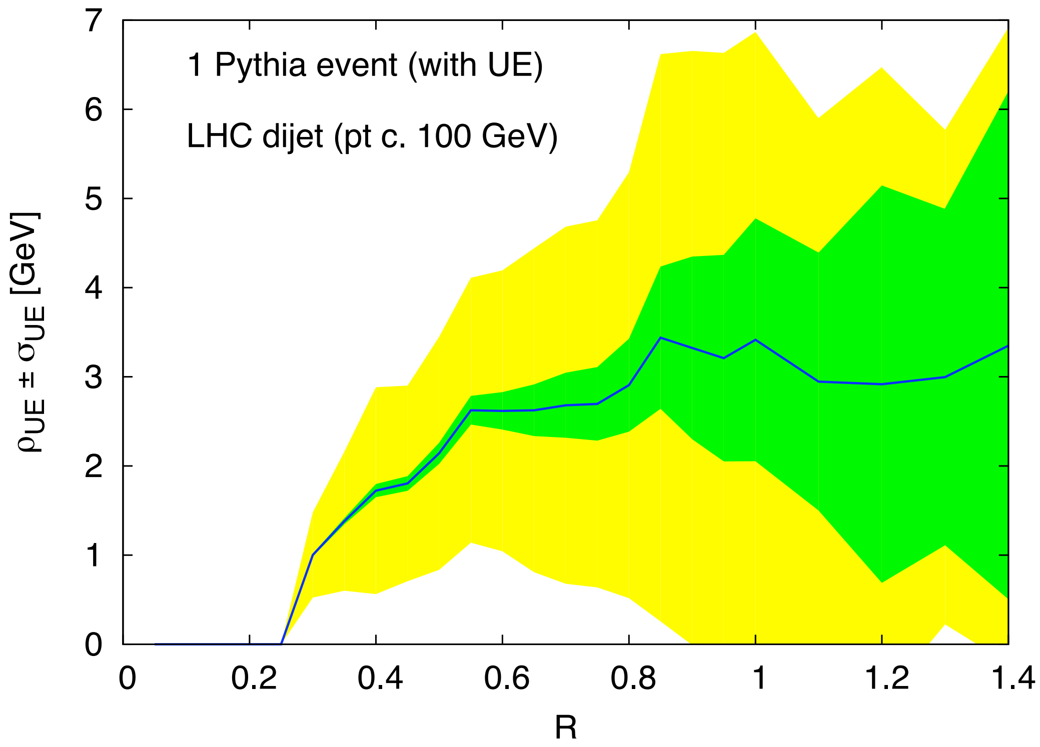


Noise levels

Pileup @ LHC



UE @ LHC



$$\rho \simeq 25\,\text{GeV}$$

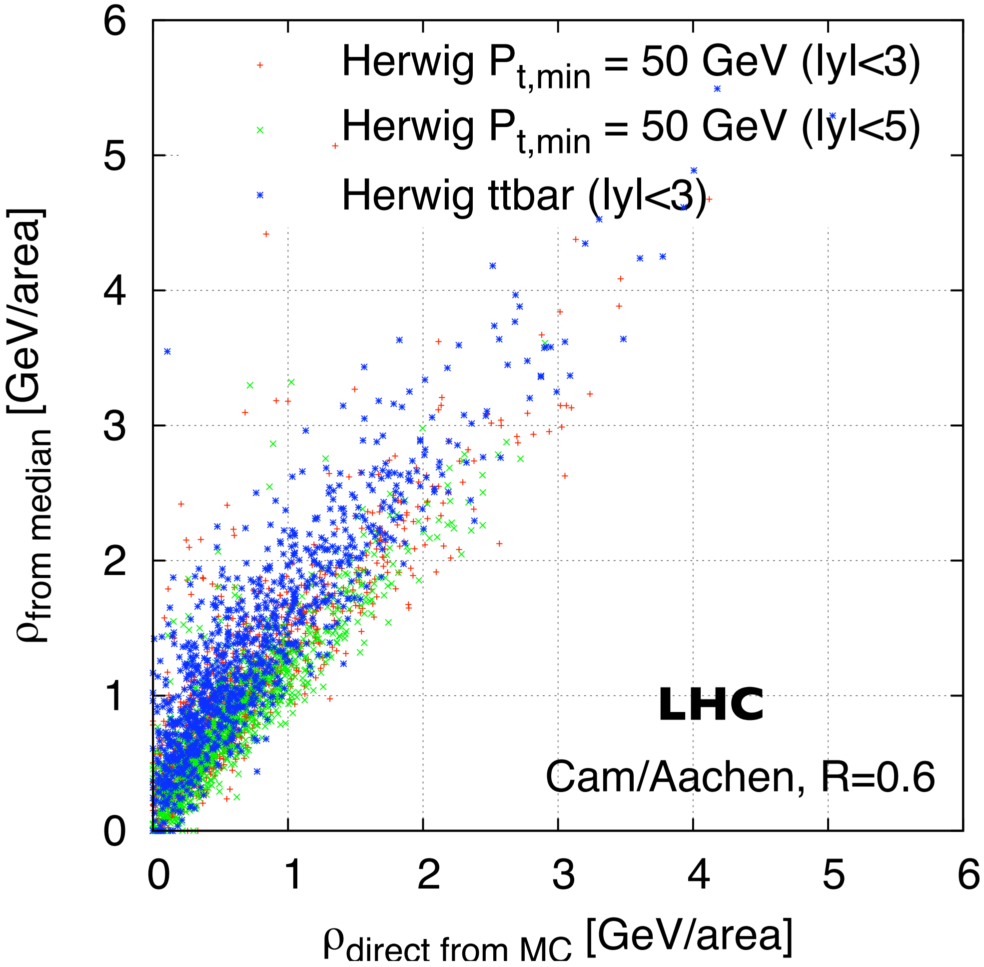
$$\rho \simeq 3\,\text{GeV}$$

Underlying Event estimation

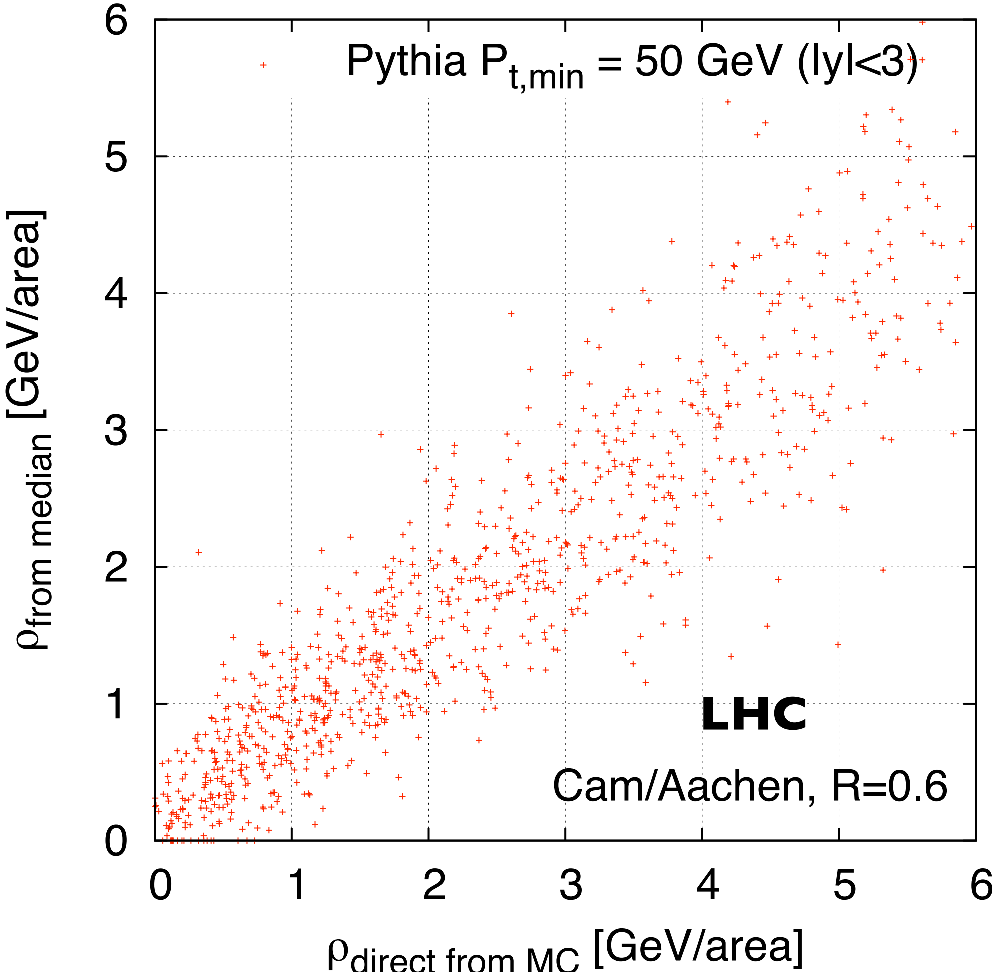
To test the procedure for the Underlying Event, compare the measurement of the background level made with areas with the known amount a Monte Carlo put in

Measurement

HERWIG

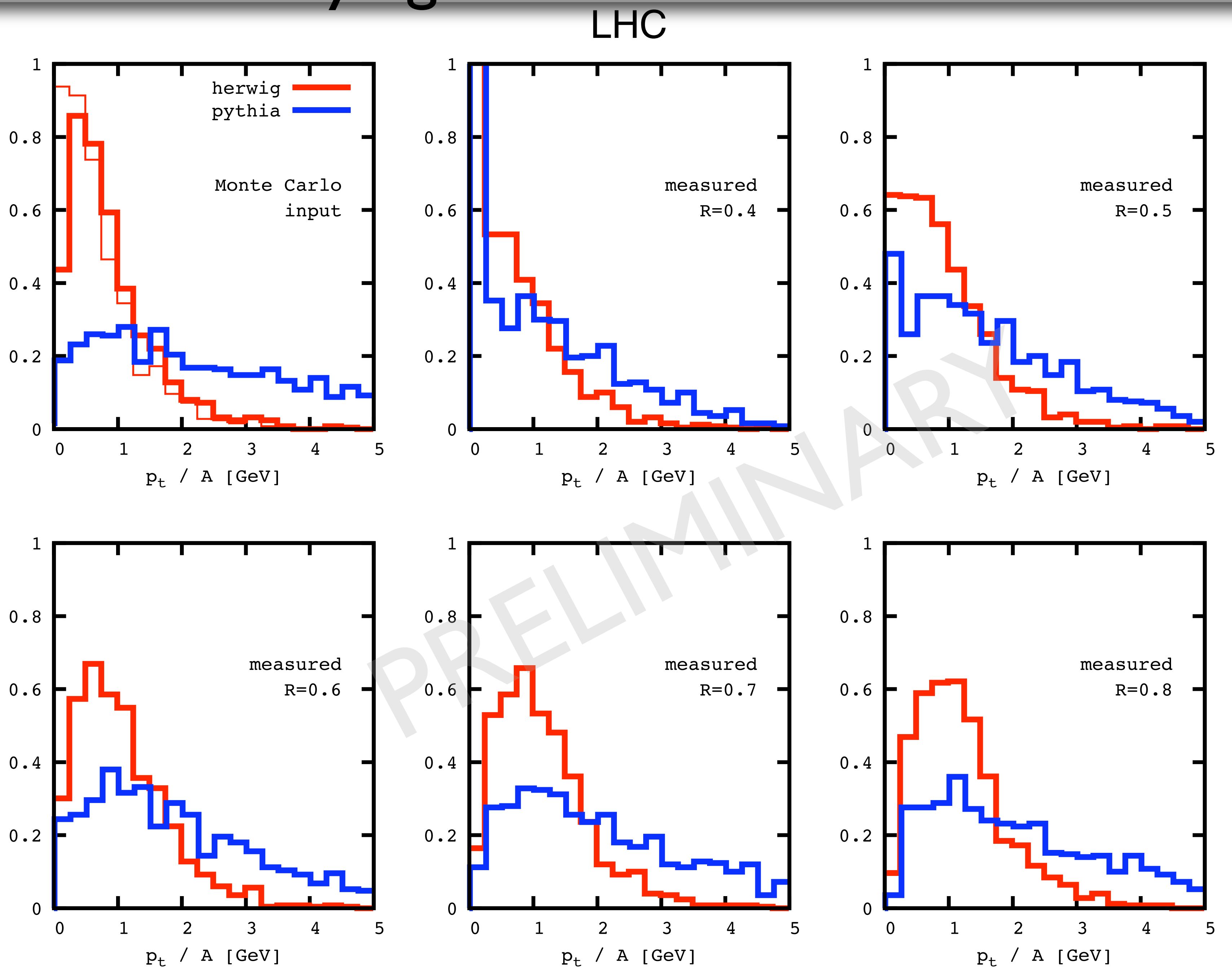


PYTHIA



Input from Monte Carlo

Underlying Event estimation: LHC



Herwig and Pythia differ. A similar analysis on the data would immediately tell which one (if either) is right

A practical application of areas: subtraction

[MC, Salam, arXiv:0707.1378]

- When a hard event is superimposed on a **roughly uniformly distributed background**, study of **transverse momentum/area** of each jet allows one to determine the noise density **ρ** (and its fluctuation) on an event-by-event basis

- Once measured, the background density can be used to correct the transverse momentum of the hard jets:

$$p_T^{\text{hard jet, corrected}} = p_T^{\text{hard jet, raw}} - \rho \times \text{Area}_{\text{hard jet}}$$

Subtraction in FastJet

```
// the input particles' 4-momenta
vector<fastjet::PseudoJet> input_particles;

// choose the jet algorithm
fastjet::JetDefinition jet_def(kt_algorithm,R);

// define the kind of area
fastjet::GhostedAreaSpec ghosted_area_spec(ghost_etamax);
fastjet::AreaDefinition area_def(ghosted_area_spec);

// perform the clustering
fastjet::ClusterSequenceArea cs(input_particles,jet_def,area_def);

// get the jets with pt > 0
vector<fastjet::PseudoJet> jets = cs.inclusive_jets();

// a jet transverse momentum, area, and area 4-vector
double pt = jets[0].perp();
double area = cs.area(jets[0]);
fastjet::Pseudojet area_4vector = cs.area_4vector(jets[0]);
```

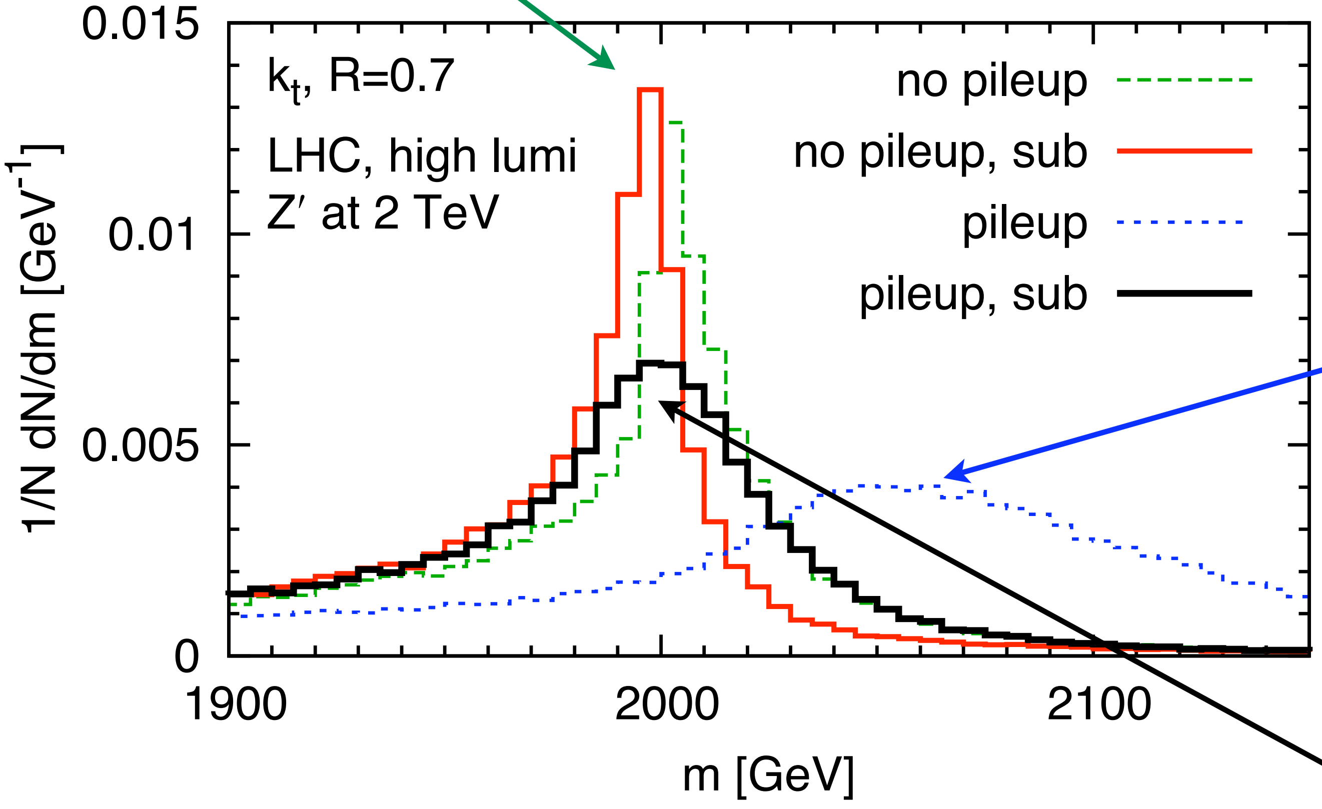
```
// get the median, i.e. rho
double rho = cs.median_pt_per_unit_area(rapmax);
double rho_4v = cs.median_pt_per_unit_area_4vector(rapmax);

// subtract
double pt_sub = pt - rho * area;
fastjet::Pseudojet p_sub = jets[0] - rho_4v * area_4vector;
```

NB. The "_4vector" variants also correct jet directions, and are better for large R

Reconstructed Z’ mass

Let’s discover a leptophobic Z’ and measure its mass:

MC simulation:
m = 2000 GeV, width ~ 10 GeV



Naive measurement with PU:
m ~ 2050 GeV, width ~ 60 GeV

Measurement after subtraction:
m ~ 2000 GeV, width ~ 25 GeV

Heavy Ion Collisions: PbPb @ LHC

Background much larger than even LHC hi-lumi pileup:

HYDJET v1.1

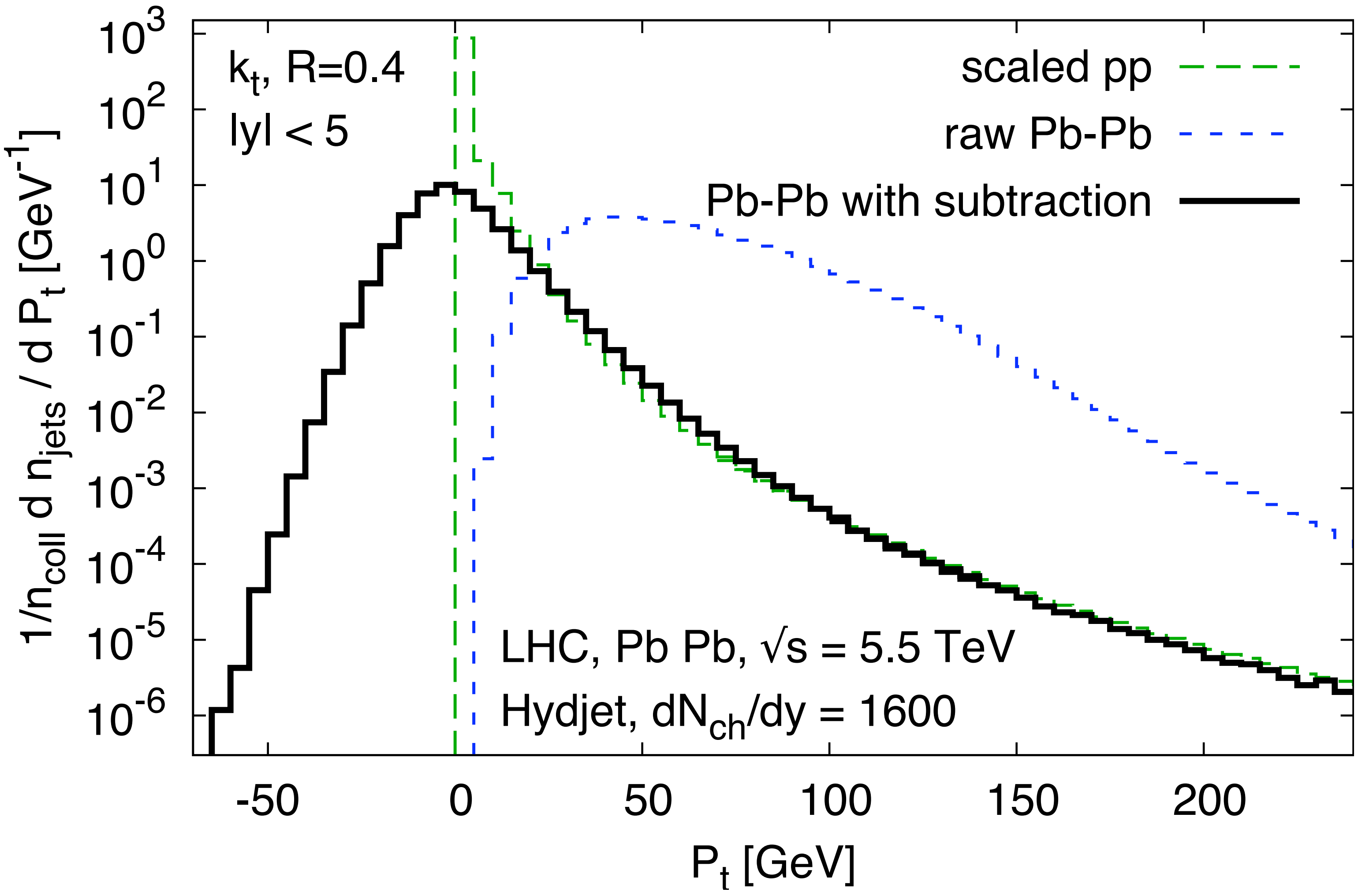
$$\left.\frac{dN_{ch}}{dy}\right|_{y=0} = 1600 \quad \Longrightarrow \quad \rho_{background} \equiv \frac{dp_T}{dy\,d\phi} \sim 250 \text{ GeV}$$

Hence, a jet with R = 0.4 on average gets an additional

$$\Delta p_T \simeq \rho_{background}\,\pi R^2 \sim 100 \text{ GeV}$$

and yet, not so much the size of this background, but rather its **fluctuations**, are the real obstacle to its subtraction

Inclusive jets in PbPb at LHC



The scaled pp cross section is recovered after subtraction

NB. No minimum pt cut
No a posteriori Monte Carlo correction

Conclusions

- Using infrared safe jet algorithms allows one to analyse them as legitimate observables in pQCD, including more exotic (and previously unexplored) characteristics like their **area**
- The area itself can be used for **background (UE and/or min-bias) estimation and subtraction**, opening the way to a more accurate, and theoretically motivated, use of jet clustering in high luminosity and even heavy ions collisions environments
- All these tools available in **FastJet** (www.lpthe.jussieu.fr/~salam/fastjet)

List of relevant papers:

MC, Salam, *Dispelling the N^3 myth for the k_t jet-finder,* hep-ph/0512210
Salam, Soyez, *A Seedles infrared safe cone algorithm,* arXiv:0704.0292
MC, Salam, *Pileup subtraction using jet areas,* arXiv:0707.1378
MC, Salam, Soyez, *The catchment area of jets,* arXiv:0802:1188
MC, Salam, Soyez, *The anti-k_t jet clustering algorithm*, arXiv:0802:1189
Les Houches 2007 proceedings, arXiv:0803.0678