Jet Areas, and What They Are Good For

Matteo Cacciari

in collaboration with

Gavin Salam

LPTHE - Paris 6,7 and CNRS
NB. Running a jet-finder algorithm gives a well defined physical observable

It’s not just some code you run to find a direction of emission of sum stuff, then patch it up for how much energy goes into it using some Monte Carlo, etc. etc.
The $k_t$ and the Cambridge/Aachen jet-finders

The longitudinally invariant $k_t$ jet-finder

Calculate the distances between the particles: $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta y^2 + \Delta \phi^2}{R^2}$

Calculate the beam distances: $d_{iB} = k_{ti}^2$

Combine particles with smallest distance or, if $d_{iB}$ is smallest, call it a jet

Find again smallest distance and repeat procedure until no particles are left

The longitudinally invariant Cambridge/Aachen jet-finder

Same as $k_t$, but with distances replaced by

$$d_{ij} = \frac{\Delta y^2 + \Delta \phi^2}{R^2} \quad d_{iB} = 1$$
A simple event

$p_T$ (parton)
The parton radiates, but we can usually collect most of its momentum into a jet.
\[ p_T \text{ (jet)} \sim p_T \text{ (parton)} + \]

Average underlying momentum density \times

‘size’ of the jet

Can we get to know the momentum density of the radiation?
Can we subtract it from the jet to find the parton momentum?
Consider an event made up of a number of particles \( \{p_i\} \), 4-momenta, calorimeter towers, ....

The 'size' of a jet can be defined by a jet-finder algorithm which extracts jets \( \{j_k\} \) from the event.
What is the ‘size’ of a jet?

The clustering procedure assigns each particle to a jet:

But... where exactly does a jet end, and another begins?
Jet Area

One idea: tile the plane, count the cells of a jet, sum the areas.

But what do I do when different jets share a cell?
Jet Area

Obviously, make the cells smaller to improve accuracy.

Unfortunately, particles being pointlike, the area tends to zero!
Next try, use the **convex hull**

But what do I do if they overlap?
Moreover, what about the ‘no man’s land’?
The Active Jet Area

We propose the following definition:

The ‘active area’ of a jet is (proportional to) the number of uniformly distributed infinitely soft particles that get clustered in it.
The Active Jet Area

After the clustering, a given set of ghosts belong to each jet

Their number (times the average area of a single ghost) defines the catchment area of the jet
The Active Jet Area

The definition of **active area** mimics the behaviour of the jet-clustering algorithms in the presence of a large number of randomly distributed soft particles

Tools needed to implement it:

1. An **infrared safe jet-finder** (the ghosts should not change the jets)
2. A reasonably **fast implementation** (we are adding thousands of ghosts) $[O(10^4)]$

Both these characteristics are found in kt and Cambridge/Aachen jet-finders (as implemented in FastJet) and in SISCones $[\sim 0.1 \text{ s}]$ $[\sim 100 \text{ s}]$

www.lpthe.jussieu.fr/~salam/fastjet
projects.hepforge.org/siscone
Purely soft jets tend to be smaller than jets ‘anchored’ by an infinitely harder particle:

\[ \langle A_{\text{soft}} \rangle \simeq 0.55 \pi R^2 \]
\[ \langle A_{\text{hard}} \rangle \simeq 0.81 \pi R^2 \]
A concrete example: a 50 GeV di-jet event at the LHC with pile-up (10 min-bias events added)
Areas distribution

The jets adapt to the surrounding environment

They can have very different areas
Area vs. $p_T$  

Key observation:  

$p_T/\text{Area}$ is fairly constant, except for the hard jets  

The distribution of background jets establishes its own average momentum density  

(NB. this is true on an event-by-event basis)
What we have seen so far

A proper operative definition of **jet area** can be given

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 $\rho$ (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:

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

A few examples follow...
Does it work? A toy model

Consider a uniform distribution of soft particles, e.g. 10000 in the rapidity range [-4,4], with $p_T = 1$ GeV.

In addition, insert a single 100 GeV particle.
Does it work? A toy model

Consider a uniform distribution of soft particles, e.g. 10000 in the rapidity range [-4,4], with $p_T = 1$ GeV

In addition, insert a single 100 GeV particle

Problem:

What will be the transverse momentum of a Cambridge/Aachen jet with $R=1$, containing the hard particle but also scooping up the background?
Does it work? A toy model

Solution:

Soft momentum density = \[ \frac{d p_T^{soft}}{dyd\phi} \equiv \rho = \frac{10000 \times 1 \text{ GeV}}{8 \times 2\pi} \approx 200 \text{ GeV} \]

Of course this number can, and in real life will, be estimated via the p_T/area median.
Solution:

Soft momentum density = \( \frac{d p_T^{soft}}{dyd\phi} \equiv \rho = \frac{10000 \times 1 \text{ GeV}}{8 \times 2\pi} \approx 200 \text{ GeV} \)

Average area of hard jet = \( \langle A_{hard} \rangle \approx \langle A_{soft} \rangle = 0.55 \times \pi R^2 \approx 1.73 \)

Of course this number can, and in real life will, be estimated via the \( p_T/\text{area median} \)

Because hard particle here not infinitely harder than soft ones
Does it work? A toy model

Solution:

Soft momentum density = \( \frac{dp_T^{soft}}{dyd\phi} \equiv \rho = \frac{10000 \times 1 \text{ GeV}}{8 \times 2\pi} \simeq 200 \text{ GeV} \)

Of course this number can, and in real life will, be estimated via the \( p_T/\text{area median} \)

Because hard particle here not infinitely harder than soft ones

Average area of hard jet = \( \langle A_{hard} \rangle \simeq \langle A_{soft} \rangle \equiv 0.55 \times \pi R^2 \simeq 1.73 \)

Typical additional soft contribution to hard jet momentum = \( \rho \times \langle A_{hard} \rangle \simeq 200 \text{ GeV} \times 1.73 \simeq 350 \text{ GeV} \)
Does it work? A toy model

The hard particle
Does it work? A toy model

The hard particle clustered with the soft background

Could just shift by estimated amount, but the resolution is hopelessly bad
Does it work? A toy model

The hard particle clustered with the soft background, **after the subtraction**

![Graph showing particle distribution](image)

The correct transverse momentum is recovered, with an important gain in resolution

NB. No cuts, no Monte Carlo correction: exclusively data driven
In decreasing order of number of particles/uniformity, we have, at the LHC,

- **Background** in heavy ion collisions
  (~ 30000 particles / event)

- **Pile-up** in high luminosity pp collisions
  (up to ~ 20 overlapping collisions, ⇒ ~ 4000 particles/event)

- **Underlying event** in a single pp collision
  (about 200 particles)

Since the measurement of the background level relies on a uniform distribution of the ‘background particles’ themselves, and assumes the background to be uncorrelated with the hard jets, we must expect the underlying event case to be the most challenging one.
Inclusive jet distribution in HIC

The momentum density of simulated events is measured to be \( \sim 250 \) GeV per unit area. Hence, with \( R = 0.4 \) a jet on average gets \( \sim 125 \) GeV of additional transverse momentum.

The jet distribution is completely distorted by the huge background.....
Inclusive jet distribution in HIC

The momentum density of simulated events is measured to be $\sim 250$ GeV per unit area. Hence, with $R = 0.4$ a jet on average gets $\sim 125$ GeV of additional transverse momentum.

The jet distribution is completely distorted by the huge background......

...but it can be recovered down to fairly low $p_T$
The peak is shifted and smeared when clustering together with the pile-up
An hypothetical $Z'$ invariant mass distribution

The correct mass is recovered, with good resolution, after subtraction
The top and $W$ mass distributions get shifted, but they can be recovered after correction with good resolution.
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.

**Input from Monte Carlo**
Given a proper jet-finder, jet areas can be defined
Conclusions

- Given a proper jet-finder, jet areas can be defined.
- They can be used to estimate the level of a uniformly distributed noise.
Conclusions

Given a proper jet-finder, jet areas can be defined.

They can be used to estimate the level of a uniformly distributed noise.

They can be used to subtract the background contribution from the hard jets. Everything is data driven: no cuts, no Monte Carlo corrections.
Conclusions

Given a proper jet-finder, jet areas can be defined

They can be used to estimate the level of a uniformly distributed noise

They can be used to subtract the background contribution from the hard jets. Everything is data driven: no cuts, no Monte Carlo corrections

Preliminary Monte Carlo tests look promising. Full ‘experimental’ tests are now needed

Work in progress
To be published soon