25th Indian-Summer School of Physics Understanding Hot & Dense QCD Matter September 2013, Prague



Matteo Cacciari LPTHE Paris

Outline

Lecture I: jet clustering

- Basic requirements and properties
- Jet clustering algorithms

Lecture 2: jet properties

- Jet areas
- Corrections to a jet transverse momentum
 - non-perturbative effects, UE, pileup,...
 - background estimation and subtraction
- Generalised pileup subtraction

Lecture 3: jet substructure

- Jet grooming
- Boosted objects: taggers
- Analytical understanding

These lectures include many ideas and contributions by Gavin Salam and Gregory Soyez

Matteo Cacciari - LPTHE

Jets 'reach'

Algorithmically, a jet is simply a collection of particles

For a number of reasons, it is however useful to consider its **spatial extent**, i.e. given the position of its axis, up to where does it collect particles? What is its shape?

These details are important for a number of corrections of various origin: perturbative, non-perturbative (hadronisation), pileup, detector related, etc

Note that the intuitive picture of a jet being a cone (of radius R) is sometimes **wrong.** This is what k_t jets can look like: (more later about what this plot really means)



Matteo Cacciari - LPTHE

From jet 'reach' to jet areas

MC, Salam, Soyez, 0802. I 188

Not one, but three **<u>definitions</u>** of a jet's size:

Passive area

Place a single very soft particle (a '**ghost**') in the event, measure the extent of the region where it gets clustered within a given jet

Reach of jet for **pointlike** radiation

Active area

Fill the events with many very soft particles ('**ghosts**'), cluster them together with the hard ones, see how many get clustered within a given jet

Reach of jet for **diffuse** radiation

• Voronoi area

Sum of areas of intersections of Voronoi cells of jet constituents with circle of radius R centred on each constituent

Coincides with passive area for k_t algorithm

(In the large number of particles limit all areas converge to the same value)

Matteo Cacciari - LPTHE

Jet active area

Active Area

Add **many** ghost particles in random configurations to the event. Cluster many times. Allow ghosts to cluster among themselves too. Count how many ghosts <u>on average</u> get clustered into a given jet J.



$$A(J) = \lim_{v_g \to \infty} \langle A(J | \{g_i\}) \rangle_g$$

Active area of jet J

Matteo Cacciari - LPTHE

Jet active 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**, like those due to **pileup or underlying event**

Tools needed to implement it

- I. 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

As a bonus, active areas also allow for a visualisation of a jet's reach

Matteo Cacciari - LPTHE









Jet areas: the single hard particle case

It is worth noting that, for a jet made of a single hard particle, while passive areas are indeed πR^2 , active areas are not

Active areas	kt	Cam/Aa	SISCone	anti-k _t
<a>/πR²	0.81	0.81	I/4	I

Only anti-k_t has the behaviour one would naively expect, i.e. area = πR^2

Active area distributions



For a roughly uniformly soft background, anti-k_t gives many small jets and many large ones (you can't fill a plane with circles!)

Matteo Cacciari - LPTHE

Areas as a dynamical jet property

The average area of a jet can change with its p_t :

$$\langle \Delta A \rangle = \mathbf{D} \frac{C_1}{\pi b_0} \ln \frac{\alpha_s(Q_0)}{\alpha_s(Rp_{t1})}$$

	k _t	Cam/Aa	SISCone	anti-k _t
D	0.52	0.08	0.12	0

Again, only anti- k_t has a typical area that does **not** increase with p_t

Matteo Cacciari - LPTHE

Jet areas scaling violations



Matteo Cacciari - LPTHE

Averages and dispersions evolution from Monte Carlo simulations (dijet events at LHC) in good agreement with simple LL calculations

Area scaling violations are a legitimate observable.

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

Understanding Hot & Dense QCD Matter - September 2013 - Prague

Jet areas scaling violations

MC, Salam, Soyez, arXiv:0802.1189

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



Jet area: summary

- ▶ Jets CAN have an area, but one must define it
- Different jet algorithms can have very different area properties:
 - Jet areas in many algorithms can fluctuate significantly from a jet to another. Isolated hard jets in anti-kt are an exception
 - Jet areas can depend on a jet's pt, driven by a (calculable) anomalous dimension that is specific to each jet algorithm. Anti-kt jets are again an exception: the anomalous dimension is zero

Jets' pt

What contributes to a jet's transverse momentum?



Contributions to a jet pt

- Ideally, we'd like a jet to contain all radiation emitted by a parton, i.e. reproduce exactly the parton. This is impossible for a number of reasons:
 - First and foremost, a parton is not a physical observable, while a jet is. There cannot be exact equivalence between them
 - A jet has finite extent, and part of the radiation (perturbative or nonperturbative) emitted by the parton may end up beyond the jet's boundaries, leading to a decrease of the jet momentum (with respect to the parton's)
 - A jet does not fragment in a vacuum. Background radiation (underlying event and/or pileup) can affect its momentum in at least two ways:
 - Some background radiation will be clustered with the jet, increasing its momentum
 - Some hard particles (from the parton fragmentation process) may not cluster with the rest of the jet because of the disturbing presence of other nearby particles (they will form other jets with them). Alternatively, hard particles that in the vacuum would not have clustered with the jet will

Effects of jet 'radius'



Irrelevant for a single-particle jet

Matteo Cacciari - LPTHE

Effects of jet 'radius'

Small jet radius



Small jet radius





Large jet radius



Large jet radius





perturbative radiation: large radius better (lose less)

non-perturbative hadronisation: large radius better (lose less)

underlying event: large radius worse (capture more)

Matteo Cacciari - LPTHE

R-dependent effects

Perturbative radiation: $\Delta p_t \simeq \frac{\alpha_s(C_F, C_A)}{\pi} p_t \ln R$

Hadronisation:
$$\Delta p_t \simeq -\frac{(C_F, C_A)}{R} \times 0.4 \text{ GeV}$$

Underlying Event:
$$\Delta p_t \simeq \frac{R^2}{2} \times (2.5 - 15 \text{ GeV})$$

Tevatron LHC

(small-R limit results) Analytical estimates: Dasgupta, Magnea, Salam, arXiv:0712.3014

Which R to choose?

The value of R matters because it affects, in opposite ways, a number of things:

Limit underlying event and pileup contamination

Small R:

Better resolve many-jets events
 Perturbative radiation loss, larger hadronisation effects

Large R: Limit non-perturbative hadronisation effects Limit perturbative radiation loss More sensitive to EU and pileup

The best compromise will in general depend on the specific observable



Minimize $\Sigma(\Delta p_t)^2$



Matteo Cacciari - LPTHE

Effect of background

How are the hard jets modified by the background?

(Can be underlying event and/or pileup)

Susceptibility (how much bkgd gets picked up)

Jet areas

Resiliency (how much the original jet changes)

Backreaction

Jet areas physical meaning

A jet's active area expresses the susceptibility of that jet's transverse momentum to contamination from a uniform background

Consider a jet of transverse momentum p_t , made up of N_g ghosts, each with transverse momentum $p_{t,g}$.

It holds $\frac{\partial p_t}{\partial p_{t,g}} = \frac{\partial (N_g p_{t,g})}{\partial p_{t,g}} = N_g$

Recalling the definition of active jet area, $A_J = A_g N_g$, we can then rewrite



The jet area is therefore the susceptibility of a jet's p_t to contamination, because for a generic background transverse momentum density ρ it will hold

$$\Delta p_t = \rho \frac{\partial p_t}{\partial \rho} = \rho A_J$$
 Susceptibility

Matteo Cacciari - LPTHE

"How (much) a jet changes when immersed in a background"

Without background



"How (much) a jet changes when immersed in a background"

Without background





"How (much) a jet changes when immersed in a background"

Without background

With background





"How (much) a jet changes when immersed in a background"

Without background

With background





Anti-kt jets are much more resilient to changes from background immersion

(NB. Backreaction is a minimal issue in pp background and at large pt. Can be much more important in Heavy Ion collisions)

Matteo Cacciari - LPTHE

The IRC safe algorithms

	Speed	Regularity	UE contamination	Backreaction	Hierarchical substructure
k t	0000	Ţ	\mathbf{T}		
Cambridge /Aachen	0000	Ţ	Ţ		
anti-k _t	0000	000	●→ ⓒ ⓒ	00	
SISCone	÷	•	☺ ☺	•	

Underlying event measurement



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**

Area v. рт

LHC: dijet event + high-lumi pilup

a few hard particles and many softer ones

(a similar picture applies to the Underlying Event)



The key observation

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



Background determination

Jet algorithms like kt or Cambridge/Aachen allow one to determine on an event-by-event basis

the "typical" level of transverse momentum density of a roughly uniform background noise:



MC, Salam, 2007

This ρ value can, in turn, be used to characterise the UE

Since this measurement is done with the jets, it is alternative/complementary to the usual analyses done using charged tracks (à la R. Field)

Background determination

MC, Salam, Sapeta, 2009

How 'right' is ρ extracted with the area/median method?

Check with a toy model of UE + two hard particles



Several features to be noted

- Too small an R underestimates ρ (too many 'empty' jets, median biased by them)
- Too large an R overestimates ρ (too few jets, median biased by hard ones)
- A positive bias from hard jets and a negative ones from soft ones tend to compensate around R = 0.5

A higher particle density (here 5 per unit area) would shift to the left the 'sweet spot' for R (and make the extraction more robust: **the more soft particles, the better it works**)

ρ can be extracted with good precision, provided an appropriate value of R is used

 Local variations of ρ (rapidity-azimuth dependent effects due to detector and/or physics) can (and must) be accounted for by a more local determination and/or a rescaling of the extracted value of ρ

Matteo Cacciari - LPTHE

UE or pileup characterisation

- ρ : average pt per unit area in an event
- $\bullet \sigma$: intra-event fluctuations
- σ_{ρ} : inter-event fluctuations



Orders of magnitude

Typical (and rough) values of ρ and σ at the LHC

	ρ	σ
pp UE	2-3 GeV	I GeV
pileup / #	I GeV	
PbPb (5.5 TeV, 0-10%)	200 GeV	20 GeV
	200 GeV	20 GeV

For an anti-kt jet of radius R=0.4, this translates into a background contamination of (100 ± 14) GeV

The HI background

Usefully characterized by its transverse momentum per unit area (ρ) and its fluctuations in a single event (σ)



ρ from ALICE data (LHC 2.76 TeV, charged only)


Hard jets and background

Modifications of the hard jet



Background subtraction

Once the **background momentum density** ρ has been measured, it can be used to **correct** the transverse momentum of the hard jets:

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

MC, Salam, 0707.1378

If ρ is measured on an event-by-event basis, and each jet subtracted individually, this procedure will remove many fluctuations and generally improve the resolution of, say, a mass peak

$$\Delta p_t = \rho A \pm (\sigma \sqrt{A} + \sigma_p A + \rho \sqrt{\langle A^2 \rangle - \langle A \rangle^2}) + \Delta p_t^{BR}$$

Irreducible fluctuations: uncertainty of the subtraction

Matteo Cacciari - LPTHE

Example of pileup subtraction

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



Hard jets and background





Hard jets (pp collisions)

Hard jets + background (AA collisions)

- In pp collisions, the background is a small correction. In heavy ions, it is overwhelming.
- It makes sense to consider background subtraction together with jet clustering: both are needed to reconstruct the jets
- As such, the same desiderata can apply: standard algorithms, well defined, with known behaviour, and well tested

Improvements in bkgd determination



Jet reconstruction in HI collisions

How do the different clustering algorithms fare?

Subtract with

$$p_{\mu,jet}^{sub} \equiv p_{\mu,jet} - \rho A_{\mu,jet}$$

Measure quality of reconstruction looking at

Offset
$$\langle \Delta p_t \rangle \equiv \langle p_t^{AA,sub} - p_t^{pp,sub} \rangle$$

Dispersion $\sigma_{jet} \equiv \sigma_{\Delta p_t} \equiv \sqrt{\langle \Delta p_t^2 \rangle - \langle \Delta p_t \rangle^2}$

[In the following I will use our own study (MC, Rojo, Salam, Soyez, 1010.1759) as a source of plots, but the results should be quite generic.
 NB. 'LHC' will be 5.5 TeV, but the results will be qualitatively similar at 2.76 TeV]

Matteo Cacciari - LPTHE

Δp_t distributions in PbPb at LHC



Matteo Cacciari - LPTHE

Understanding Hot & Dense QCD Matter - September 2013 - Prague







Dispersion of $\Delta p_t = \sigma_{\Delta pt}$



- C/A(filt) markedly better, as a consequence of its smaller effective area
- Dispersions increase at large $p_t, as a \mbox{ consequence of a larger dispersion of back-reaction}$
- anti-k_t remains fairly constant ('resiliency'), and eventually becomes better at large p_t

Matteo Cacciari - LPTHE

Jet shapes

A jet shape is any function of the momenta of a jet's constituents

- The simplest example is the **scalar transverse momentum**: sum of the pt's of all constituents
- Other IRC-safe examples: the jet's mass, splitting scales, N-subjettiness, energy energy correlation,....
- Non-IRC-safe examples are the moments of a jet fragmentation function
- ▶ Jet shapes can be used to characterise the jets
- Even better, if calculable they can be compared to theoretical predictions

However, they will in general be affected by background. How to subtract its contamination?

More on pileup subtraction

The **p_T^{raw}-ρA** technique (also called **area/median**) only corrects a jet's transverse momentum

Each jet shape has its own specific sensitivity to background contamination. How to correct them?

- One option is to study analytically each shape [Alon et al. 1101.3002]. Can be time consuming and cumbersome
- Alternatively, determine numerically the susceptibility of any IRC-safe jet shape to contamination [Soyez et al. 1211.2811] (this generalises the jet area)

Numerical jet shape correction



Numerical jet shape correction



Numerical jet shape correction Example: T₃₂ correction and top tagging



 Original distribution reproduced after pileup subtraction



 Tagging rates independent of amount of pileup after correction of the jet shapes involved in the tagging

subtraction

Jet shape subtraction in ATLAS



Matteo Cacciari - LPTHE

ATLAS-CONF-2013-085

Summary

- The susceptibility of a jet's transverse momentum can be determined in terms of a jet area (or, for a generic jet shape, a numerical derivative with respect to ghosts momenta)
- Background densities can be estimated with a number of techniques, some of them jet-based. Non uniformities in the background can, up to a certain extent, be accounted for.
- The background can then be subtracted from jet momenta or from jet-shapes
 - Alternative subtraction techniques can be devised. Usually, there will be a compromise between
 residual bias and dispersion
- All these tools and techniques are implemented in a standard way in Fastjet or its 3rd-party add-on fjcontrib (fastjet.hepforge.org)

<Areas anomalous dimensions>

Jet areas: a second hard(ish) particle

 Δ_{12}

Real events have more than a single hard particle. Add a second (soft) one at a distance Δ_{12}



Passive areas (and SISCone's active area) can be calculated analytically, while the others are obtained numerically

Matteo Cacciari - LPTHE

Jet areas: anomalous dimensions

Finally, weigh the probability of emission of the soft particle with the leading-order QCD matrix element:

$$\langle \Delta area \rangle = \int C_1 \frac{\alpha_s(p_{t2}\Delta_{12})}{\pi} \frac{dp_{t2}}{p_{t2}} \left[\frac{d\Delta_{12}}{\Delta_{12}} \right]_+ \begin{pmatrix} 1 & \Delta_{12} & 2\\ \bullet & --- & \bullet\\ hard & soft \end{pmatrix}$$

The result is an **anomalous dimension**.

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

$$\langle \Delta area
angle = \mathbf{d} \; \; rac{C_1}{\pi b_0} \ln rac{lpha_s(Q_0)}{lpha_s(Rp_{t1})}$$

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

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

Matteo Cacciari - LPTHE

Passive areas: analytical results

MC, Salam, Soyez, arXiv:0802.1188

$$\begin{aligned} d_{k_{t},R} &= \left(\frac{\sqrt{3}}{8} + \frac{\pi}{3} + \xi\right) R^{2} \simeq 0.5638 \,\pi R^{2} \,, \\ d_{\text{Cam},R} &= \left(\frac{\sqrt{3}}{8} + \frac{\pi}{3} - 2\xi\right) R^{2} \simeq 0.07918 \,\pi R^{2} \,, \\ d_{\text{SISCone},R} &= \left(-\frac{\sqrt{3}}{8} + \frac{\pi}{6} - \xi\right) R^{2} \simeq -0.06378 \,\pi R^{2} \,, \\ \begin{array}{l} \text{Negative!} \\ \text{SISCone} \\ \text{jets shrink!} \end{array}$$

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

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

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

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

Matteo Cacciari - LPTHE

d:

Jet areas: passive v. active

	area/ π R ²		dispersion		d or D		s or S				
	passive	active	passive	active	passive	active	passive	active			
	$a(1\mathrm{PJ})$	$A(1\mathrm{PJ})$	$\sigma(1PJ)$	$\Sigma(1\mathrm{PJ})$	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			
	sii	emission of a second perturbative particle (coeff. of anomalous dimension)									

Jet areas: passive v. active

	area/ π R ²		dispersion		d or D		s or S				
	passive	active	passive	active	passive	active	passive	active			
	$a(1\mathrm{PJ})$	$A(1\mathrm{PJ})$	$\sigma(1PJ)$	$\Sigma(1\mathrm{PJ})$	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			
	siı	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-k_t has constant area (null anomalous dimension): it's a **perfect cone**

Matteo Cacciari - LPTHE

</Areas anomalous dimensions>